

Accession # 100453

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Name Moshur Accession # 67746 Date 8/5
 Accession # 1043 Phone Number 308-2926 Serial Number 091851,410
 Mail Box and Bag Room Location _____ Results Format Preferred (circle) PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Provide a detailed description of the search topic and describe in detail the information and subject matter to be searched. Include the date and time of the search, keywords, synonyms, acronyms, and any other numbers and combine with the concept or subject of the invention. List the key terms that may have a special meaning. Give examples of relevant citations, authors, etc., if known. Please attach a copy of the patent, pertinent claims, and abstract.

Title of Invention _____

Inventor(s) (provide full names) _____

Earliest Priority Filing Date _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Search seq 8 str & oligo

STAFF USE ONLY

STAFF USE ONLY		Type of Search	Vendors and cost where applicable
Searcher <u>D. Schweitzer</u>	NA Sequence # _____	CM	_____
Searcher Phone # <u>308-4292</u>	AA Sequence # <u>2</u>	Dialog	_____
Searcher Location <u>CM1 6A03</u>	Structure # _____	Desktop	_____
Searcher Initials <u>DS</u>	Biographic _____	_____	_____
Searcher Date <u>8/5</u>	Ligation _____	_____	_____
Searcher ID & Ref. # <u>13</u>	Full text _____	Sequence Analysis	<u>Com pay</u>
Searcher Dept. <u>6</u>	Parent Name _____	_____	_____
Searcher Initials <u>6</u>	Other _____	_____	_____

The first part of the paper discusses the importance of the study and the objectives of the research. It highlights the need for a comprehensive understanding of the subject matter and the role of the researcher in this process. The second part of the paper presents the methodology used in the study, including the data collection methods and the analysis techniques. The third part of the paper discusses the results of the study and the conclusions drawn from the findings. The final part of the paper provides a summary of the key points and offers suggestions for future research.

The study was conducted in a systematic and rigorous manner, following the principles of scientific research. The data was collected from a large sample of participants, and the results were analyzed using advanced statistical techniques. The findings of the study are presented in a clear and concise manner, allowing for a thorough understanding of the subject matter. The conclusions drawn from the findings are based on a careful analysis of the data and are supported by the results of the study.

The study has several limitations, which are discussed in the paper. These limitations include the sample size, the duration of the study, and the potential for bias. Despite these limitations, the study provides valuable insights into the subject matter and contributes to the existing body of knowledge. The findings of the study are discussed in the context of the current research and are compared with the results of previous studies.

The study has several implications for practice and policy. The findings of the study suggest that there is a need for further research in this area, and that the results of the study can be used to inform decision-making. The study also highlights the importance of the role of the researcher in this process and the need for a comprehensive understanding of the subject matter.

In conclusion, the study provides a comprehensive understanding of the subject matter and contributes to the existing body of knowledge. The findings of the study are presented in a clear and concise manner, allowing for a thorough understanding of the subject matter. The conclusions drawn from the findings are based on a careful analysis of the data and are supported by the results of the study.



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number

TO: Mary Mosher
Location: cm-1/9a17/8e12
Art Unit: 1648
Tuesday, August 05, 2003

Case Serial Number: 09/851410

From: David Schreiber
Location: Biotech-Chem Library
CM1-6A03
Phone: 308-4292

david.schreiber@uspto.gov

Search Notes

Search seq 8 std and oligo

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM Protein protein search, using sw model

Run on: August 5, 2003, 09:29:48 : Search time 47 Seconds
(without alignments)
2228,926 Million cell updates/sec

File: US-09-851-410A.8

Peptide score: 660

Sequence: 1 MRRPRLIIIIITIMEIPMPAP.....QSTVAQGLPEWVETPEL 660

Scoring table:

Gapop: 60.0 ; Gapext: 60.0

Search-E: 1107863 seqs, 158726574 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1107863

Minimum db seq length: 0

Maximum db seq length: 2900000000

Post processing: listing first 45 summaries

Database: A_Genescq_19Jun03.*

```

1: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1980.DAT.*
2: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1981.DAT.*
3: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1982.DAT.*
4: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1983.DAT.*
5: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1984.DAT.*
6: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1985.DAT.*
7: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1986.DAT.*
8: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1987.DAT.*
9: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1988.DAT.*
10: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1989.DAT.*
11: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1990.DAT.*
12: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1991.DAT.*
13: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1992.DAT.*
14: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1993.DAT.*
15: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1994.DAT.*
16: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1995.DAT.*
17: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1996.DAT.*
18: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1997.DAT.*
19: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1998.DAT.*
20: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA1999.DAT.*
21: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA2000.DAT.*
22: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA2001.DAT.*
23: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA2002.DAT.*
24: ZSUS1/ncdata/a/genescq/genescsq-emb1/AA2003.DAT.*

```

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	660	100.0	660	14 AAR48785	HEV ORF2 protein.
2	660	100.0	660	14 AAR49506	Burma strain HEV O
3	660	100.0	660	17 AAR46089	Hepatitis E virus
4	660	100.0	660	18 AAR35826	Protein encoded by
5	660	100.0	660	19 AAR80197	Human HEV ORF 2 pr
6	660	100.0	660	20 AAR43389	Hepatitis E virus
7	660	100.0	660	21 AAR24120	HEV Burma strain v
8	660	100.0	660	21 AAR62523	Hepatitis E virus
9	660	100.0	660	23 AAR15699	Hepatitis E virus

10	559	84.7	660	19 AAR76369	Hepatitis E virus
11	559	84.7	660	19 AAR71210	Protein encoded by
12	549	83.2	649	17 AAR66091	Hepatitis E virus
13	541	82.0	660	22 AAR12980	Hepatitis E virus
14	540	81.8	540	17 AAR66101	Hepatitis E virus
15	525	79.5	525	17 AAR66103	Hepatitis E virus
16	512	77.6	660	19 AAR76324	Hepatitis E virus
17	448	67.9	549	19 AAR76367	Hepatitis E virus
18	446	67.6	660	23 AAR104004	Hepatitis E virus
19	438	61.8	660	23 AAR91811	Hepatitis E virus
20	408	61.8	660	20 AAR43386	Human HEV ORF 2 pr
21	408	61.8	660	24 AAR81670	Hepatitis E virus
22	396	60.0	660	20 AAR43388	Human HEV ORF 2 pr
23	327	49.5	327	17 AAR66093	Hepatitis E virus
24	327	49.5	327	18 AAR45820	Hepatitis E virus
25	321	48.6	660	15 AAR51265	HEV strain protein
26	321	48.6	660	19 AAR81520	Hepatitis E virus
27	310	47.0	660	24 AAR104816	Hepatitis E virus
28	304	46.1	660	24 AAR104820	Hepatitis E virus
29	298	45.2	660	24 AAR104811	Hepatitis E virus
30	284	44.0	660	20 AAR43395	Human HEV ORF 2 pr
31	264	40.0	660	20 AAR43392	Human HEV ORF 2 pr
32	259	39.2	660	20 AAR43390	Human HEV ORF 2 pr
33	255	38.6	660	24 AAR104807	Hepatitis E virus
34	251	38.0	660	24 AAR104809	Hepatitis E virus
35	251	38.0	660	24 AAR104818	Hepatitis E virus
36	250	37.9	660	24 AAR104819	Hepatitis E virus
37	228	34.5	660	12 AAR104619	Protein encoded by
38	214	32.4	660	24 AAR104821	Hepatitis E virus
39	214	32.4	660	24 AAR104822	Hepatitis E virus
40	210	31.6	213	24 AAR10496	Hepatitis E virus
41	210	31.6	213	24 AAR10497	Hepatitis E virus
42	210	31.6	232	24 AAR81679	Hepatitis E virus
43	200	30.3	489	24 AAR81674	Hepatitis E virus
44	197	29.8	660	20 AAR43391	Human HEV ORF 2 pr
45	196	29.7	660	24 AAR104823	Hepatitis E virus

ALL SUMMARIES

RESULT 1
AAR48785
ID: AAR48785 Standard: Protein, 660 AA
XX
AA: AAR48785;
XX
25-MAR-2003 (updated)
11-JAN-1994 (first entry)
HEV ORF2 protein.
XX
Hepatitis E virus Burma strain.
XX
Key location/Accession
FT Peptide 225..660
FT Peptide /Label 25
FT Peptide 334..660
FT Peptide /Label 334
FT Peptide 613..660
FT Peptide /Label 403, 427
XX
W09314116-AL
XX
22-JUL-1993
XX
15-JAN-1993; 9AAR0-US004
XX
17-JAN-1992; 9AAR0-US004
XX
01-MAY-1992; 9AAR0-US004

XX (GENE-) GENELABS TECHNOLOGIES INC.
 PA (USSH) US SEC DEPT HEALTH.
 XX Bradley DW, Carl M, Reyes GR, Tam AW,
 XX WPI: 1993-243144/30.
 DR N-PSDB: AAQ47129.
 XX
 PT New immunogenic hepatitis E virus (HEV) peptide(s) - are from the
 PI ORF1, ORF2 and ORF3 regions of HEV, useful as a vaccine against
 PT HEV infection
 XX
 PS Disclosure, Fig 8, 48pp, English.
 XX
 CC Immunogenic hepatitis E virus (HEV) peptides are selected from the
 CC ORF1, ORF2 and ORF3 regions of HEV. The peptides can be used in
 CC vaccines to prevent infection by HEV. The antibodies can neutralise
 CC and block HEV infection and can be used to prevent or treat HEV
 CC infection. The peptides and antibodies can also be used as
 CC diagnostic reagents.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 660 AA:
 Query Match 100.0%; Score 660; DB 14; Length 660;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 660, Conservative 0, Mismatches 0, Indels 0, Gaps 0;
 QY 1 MRPPPTILLLMFLPMI PAPPDQPGSPRRPSPSGSGSGGFWDFVTSPPFAIPIYTHPTN 60
 DB 1 MRPPPTILLLMFLPMI PAPPDQPGSPRRPSPSGSGSGGFWDFVTSPPFAIPIYTHPTN 60
 QY 61 PFAFVTAAGAGPRVQPPAPPIGCSAMPDQAPPAVASPRPTTAGAAPLTAVAPAHDP 120
 DB 61 PFAFVTAAGAGPRVQPPAPPIGCSAMPDQAPPAVASPRPTTAGAAPLTAVAPAHDP 120
 QY 121 PVPDVSRGAILRQYNLSLSPITSSVAIGINLVLYAAPLSPLPLQDGTNTHIMATEAS 180
 DB 121 PVPDVSRGAILRQYNLSLSPITSSVAIGINLVLYAAPLSPLPLQDGTNTHIMATEAS 180
 QY 181 NYAGYVARATIPYPLVPAVNAQGYAISTFWPQTITTTTSVDMNSITSTWPLVQPCI 240
 DB 181 NYAGYVARATIPYPLVPAVNAQGYAISTFWPQTITTTTSVDMNSITSTWPLVQPCI 240
 QY 241 ASELVTPSERIHYPNQWRVFTSGVAEFAEATSGIWMICIHGSLVNSYNTPTGALGIL 300
 DB 241 ASELVTPSERIHYPNQWRVFTSGVAEFAEATSGIWMICIHGSLVNSYNTPTGALGIL 300
 QY 301 DFALELEFPNITPGNTNTRVSPYSSSTAPHPLPFGAGTAEITTAATPEMKDLTSTNG 360
 DB 301 DFALELEFPNITPGNTNTRVSPYSSSTAPHPLPFGAGTAEITTAATPEMKDLTSTNG 360
 QY 361 VGRGRTATITLFIADTLGGLPTLISACVQVIEYSPVVSANCFEFTVKILYTSVENAQ 420
 DB 361 VGRGRTATITLFIADTLGGLPTLISACVQVIEYSPVVSANCFEFTVKILYTSVENAQ 420
 QY 421 GCKRTATITPDTLSPVVIQYVQNHIEQIPTTSFAPSPFVILKARVIMLSIAAPY 480
 DB 421 QDKGIAITPDIIDLGSRVVIQYVQNHIEQIPTTSFAPSPFVILKARVIMLSIAAPY 480
 QY 481 DQSTYGSSTSPVVSQSVTLVNVATCAQAVAPSLDWTKVTLDGPPILSTIQYSKTFPVLP 540
 DB 481 DQSTYGSSTSPVVSQSVTLVNVATCAQAVAPSLDWTKVTLDGPPILSTIQYSKTFPVLP 540
 QY 541 LPKRLSPWEACTTKAGYPYNYNTTASDOLLVENAAGHPVAISTYTTIS/CAGDPVSIASAV 600
 DB 541 LPKRLSPWEACTTKAGYPYNYNTTASDOLLVENAAGHPVAISTYTTIS/CAGDPVSIASAV 600
 QY 601 LAPHSALEEDILYPAFAHPIHGPYTPPHIGLQDAPASIVAPIQFLKMKVSKTPEL 660
 DB 601 LAPHSALEEDILYPAFAHPIHGPYTPPHIGLQDAPASIVAPIQFLKMKVSKTPEL 660

RESULT 2
 AAR39306
 ID AAR39306 standard: Protein: 660 AA.
 XX
 AC AAR39306;
 XX
 DT 25-MAR-2003 (updated)
 DT 14-FEB-1994 (first entry)
 XX
 PP Porra strain HEV ORF2 putative virus capsid protein.
 XX
 KW Hepatitis E virus, vaccine, neutralising antibodies, infection;
 KW block; open reading frame; antibodies.
 XX
 GS Hepatitis E virus.
 XX
 PW W09314208 A2.
 XX
 PD 22-JUL-1993.
 XX
 ZI 19 JAN 1993; 93WO-US00475.
 XX
 PR 17-JAN-1992; 92US-0822335.
 PP 20-APR-1992; 92US-0870985.
 XX
 PA (GENE) GENELABS TECHNOLOGIES INC.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 RA Bradley DW, Krawczynski KZ, Purdy MA, Reyes GR, Tam AW, Two J;
 FC WPI: 1993-243144/30.
 XX
 NP N-PSDB: AAQ46813.
 DR
 XX
 PT Antigen and antibody vaccines against hepatitis E virus infection
 PT - contain peptides derived from capsid protein of viral or
 PT antibodies against protein
 XX
 PS Disclosure; Fig 7; 43pp; English.
 XX
 CC The sequence is that of the putative virus capsid protein encoded
 CC by Burma strain hepatitis E virus (HEV) open reading frame ORF2.
 CC This protein or peptide fragments of it may be used in a vaccine
 CC composition for immunising an individual against HEV. Antibodies
 CC raised against these peptides can also be used in such vaccines.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 660 AA;
 Query Match 100.0%; Score 660; DB 14; Length 660;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 660, Conservative 0, Mismatches 0, Indels 0, Gaps 0;
 QY 1 MRPPPTILLLMFLPMI PAPPDQPGSPRRPSPSGSGSGGFWDFVTSPPFAIPIYTHPTN 60
 DB 1 MRPPPTILLLMFLPMI PAPPDQPGSPRRPSPSGSGSGGFWDFVTSPPFAIPIYTHPTN 60
 QY 61 PFAFVTAAGAGPRVQPPAPPIGCSAMPDQAPPAVASPRPTTAGAAPLTAVAPAHDP 120
 DB 61 PFAFVTAAGAGPRVQPPAPPIGCSAMPDQAPPAVASPRPTTAGAAPLTAVAPAHDP 120
 QY 121 PVPDVSRGAILRQYNLSLSPITSSVAIGINLVLYAAPLSPLPLQDGTNTHIMATEAS 180
 DB 121 PVPDVSRGAILRQYNLSLSPITSSVAIGINLVLYAAPLSPLPLQDGTNTHIMATEAS 180
 QY 181 NYAGYVARATIPYPLVPAVNAQGYAISTFWPQTITTTTSVDMNSITSTWPLVQPCI 240
 DB 181 NYAGYVARATIPYPLVPAVNAQGYAISTFWPQTITTTTSVDMNSITSTWPLVQPCI 240
 QY 241 ASELVTPSERIHYPNQWRVFTSGVAEFAEATSGIWMICIHGSLVNSYNTPTGALGIL 300
 DB 241 ASELVTPSERIHYPNQWRVFTSGVAEFAEATSGIWMICIHGSLVNSYNTPTGALGIL 300


```

PI Reyes JR, Yarbough JW, Kravitzynski KZ, Lam AW, Fry KL,
XX WFL: 2001 34:2705-2716.
DB N-PSDB: AAF83495.
XX
PI New DNA sequences of enterically transmitted non-A/non-B (ET-NANB)
PI hepatitis viral agent, useful in diagnosing infection by an enterically
PI transmitted agent (e.g. ET-NANB virus), as well as in vaccine
PI production.
XX
PS Disclosure: Columns 64-68: 45pp; English.
XX
XX The invention relates to an isolated DNA comprising the genome of an
XX enterically transmitted nonA/nonB (ET-NANB) viral hepatitis agent (also
XX referred as HEV). The DNA sequences or their fragments are useful in
XX preparing ET-NANB viral proteins and as probes for virus detection.
XX These are particularly useful in diagnosing infection by an enterically
XX transmitted agent (e.g. ET-NANB virus), as well as in vaccine production.
XX The present sequence represents a open? protein encoded by a ET-NANB viral
XX DNA sequence from HEV-Burma strain.
XX
SQ Sequence 650 AA:
Query Match 100.0%; Score 660; Ig 22; Length 660;
Best Local Similarity 100.0%; Pred. No: 0;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPPRPILLLLIMFPLMLAPPQVQSGRRGRSGGSGFWGDRVUSQFPAIYDPTN 60
DB 1 MPPRPILLLLIMFPLMLAPPQVQSGRRGRSGGSGFWGDRVUSQFPAIYDPTN 60
QY 61 PRAPVTAAGAGPRVQPARPLQSSAMWQAGPQAVASRRPPTTAGAAPLTAVAPADTP 120
DB 61 PRAPVTAAGAGPRVQPARPLQSSAMWQAGPQAVASRRPPTTAGAAPLTAVAPADTP 120
QY 121 PPDVLSKCAILRQYNIISPLTSSVATGTLNLYAAPLSPLLPLQDGTNTHIMAFAS 180
DB 121 PPDVLSKCAILRQYNIISPLTSSVATGTLNLYAAPLSPLLPLQDGTNTHIMAFAS 180
QY 181 NYAQYKVAKATIRYKPLVNAVGGYASISFWPQTITPTTSVDVMSITSDVRLVQPGI 240
DB 181 NYAQYKVAKATIRYKPLVNAVGGYASISFWPQTITPTTSVDVMSITSDVRLVQPGI 240
QY 241 ASSELVPSPIHYVQGWPSVFTSGVAFFEATSGVLMCLTHGSLVNSYTNTPYTGALCL 300
DB 241 ASSELVPSKILHYRQGWKSVFTSGVAFFEATSGVLMCLTHGSLVNSYTNTPYTGALCL 300
QY 301 DHALLELEPNLIPGNTNTRVSKYSTAPHPITPPADPTAFITTTAATPPKPDLYFTSTNG 360
DB 301 DHALLELEPNLIPGNTNTRVSKYSTAIRILRGADGTAEITTTAATRPKMDLYFTSTNG 360
QY 361 VKEIGGIALTLFNLADILLGDTPELISAGGQIPYSPPVVSANGPTVKLYTSVHNAQ 420
DB 361 VKEIGGIALTLFNLADILLGDTPELISAGGQIPYSPPVVSANGPTVKLYTSVHNAQ 420
QY 421 QERGLADPRIDGDESRVLTGQVQKHEQDPTTSPAFSPFVSLEAKWMLISLAIFY 480
DB 421 QERGLADPRIDGDESRVLTGQVQKHEQDPTTSPAFSPFVSLEAKWMLISLAIFY 480
QY 481 POSTYGSSTGPVVSIVLNVNAPAGAVAPSLWPKVILKCPPLSTLYSKTFEVL 540
DB 481 POSTYGSSTGPVVSIVLNVNAPAGAVAPSLWPKVILKCPPLSTLYSKTFEVL 540
QY 541 LPKRLSEWIACTEAWTLYRSTLAAAGLVNAAQHEVALGLYLSLAAQVGLSAAVAV 600
DB 541 LPKRLSEWIACTEAWTLYRSTLAAAGLVNAAQHEVALGLYLSLAAQVGLSAAVAV 600
QY 601 LAHESALAILFLLIYPARAHITFDYPTFWGLGLQVAFQSTVARELQPLKMKVGTTEL 660
DB 601 LAHESALAILFLLIYPARAHITFDYPTFWGLGLQVAFQSTVARELQPLKMKVGTTEL 660

```

RESULT 9

```

AA015699
PI AA015699 standard: Protein; 500 AA.
XX
AC AA015699;
XX
XX 08-NOV-2002 (first entry)
XX
XX Hepatitis E virus (Burma strain) ORF2 protein.
DE
XX HEV, enterically transmitted nonA/nonB hepatitis viral agent;
KW Burma strain; bile; ORF2.
XX
XX Hepatitis E virus.
OS
XX US6379891 BL.
XX
XX 30-APR 2002.
XX
PI 19 APR 2002; 200303 05:44Z.
XX
XX 25-JUL-1994; 9405-027082Z.
XX
XX 07-JUN-1995; 9405-047850Z.
XX
XX 15-APR-1991; 9105-0681078.
XX
XX 17-JUN-1989; 8805-0208997.
XX
XX 11-APR-1989; 8905-0336672.
XX
XX 16-JUN-1989; 8905-0367486.
XX
XX 13-OCT-1989; 8905-042092Z.
XX
XX 05-APR-1990; 9005-0505888.
XX
XX (USSH) US Dept Health & Human Services.
XX (GENE-) GENELABS TECHNLOGIES INC.
XX
XX Reyes JR, Yarbough JW, Kravitzynski KZ, Lam AW, Fry KL;
XX WFL: 2002-51727755.
XX N-PSDB: AAL50486.
XX
XX Detecting the presence of enterically transmitted nonA/nonB hepatitis
XX viral (HEV) agents in bile samples from infected humans and monkeys
XX using polymerase chain reaction.
XX
XX Disclosure: Column 19-24: 6pp; English.
XX
XX The invention comprises a method for detecting the presence of
XX enterically transmitted nonA/nonB hepatitis viral (HEV) agents in a
XX sample and isolating HEV agents or nucleic acid fragments produced by the
XX agent. The method utilizes PCR using bile from a human or cynomolgus
XX monkey actively infected with HEV as a source of the agent. The method of
XX the invention is used for detecting the presence of a viral agent in a
XX sample of cultured cells infected with the agent and isolating
XX enterically transmitted nonA/nonB HEV agents or nucleic acid fragments
XX produced by the agent. The present amino acid sequence represents the
XX protein encoded by open reading frame 2 (ORF2) of a DNA sequence isolated
XX from a Burmese strain of the hepatitis E virus.
XX
XX Sequence 650 AA:
Query Match 100.0%; Score 660; Ig 22; Length 660;
Best Local Similarity 100.0%; Pred. No: 0;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPPRPILLLLIMFPLMLAPPQVQSGRRGRSGGSGFWGDRVUSQFPAIYDPTN 60
DB 1 MPPRPILLLLIMFPLMLAPPQVQSGRRGRSGGSGFWGDRVUSQFPAIYDPTN 60
QY 61 PRAPVTAAGAGPRVQPARPLQSSAMWQAGPQAVASRRPPTTAGAAPLTAVAPADTP 120
DB 61 PRAPVTAAGAGPRVQPARPLQSSAMWQAGPQAVASRRPPTTAGAAPLTAVAPADTP 120
QY 121 PPDVLSKCAILRQYNIISPLTSSVATGTLNLYAAPLSPLLPLQDGTNTHIMAFAS 180
DB 121 PPDVLSKCAILRQYNIISPLTSSVATGTLNLYAAPLSPLLPLQDGTNTHIMAFAS 180
QY 181 NYAQYKVAKATIRYKPLVNAVGGYASISFWPQTITPTTSVDVMSITSDVRLVQPGI 240
DB 181 NYAQYKVAKATIRYKPLVNAVGGYASISFWPQTITPTTSVDVMSITSDVRLVQPGI 240
QY 241 ASSELVPSPIHYVQGWPSVFTSGVAFFEATSGVLMCLTHGSLVNSYTNTPYTGALCL 300
DB 241 ASSELVPSKILHYRQGWKSVFTSGVAFFEATSGVLMCLTHGSLVNSYTNTPYTGALCL 300
QY 301 DHALLELEPNLIPGNTNTRVSKYSTAPHPITPPADPTAFITTTAATPPKPDLYFTSTNG 360
DB 301 DHALLELEPNLIPGNTNTRVSKYSTAIRILRGADGTAEITTTAATRPKMDLYFTSTNG 360
QY 361 VKEIGGIALTLFNLADILLGDTPELISAGGQIPYSPPVVSANGPTVKLYTSVHNAQ 420
DB 361 VKEIGGIALTLFNLADILLGDTPELISAGGQIPYSPPVVSANGPTVKLYTSVHNAQ 420
QY 421 QERGLADPRIDGDESRVLTGQVQKHEQDPTTSPAFSPFVSLEAKWMLISLAIFY 480
DB 421 QERGLADPRIDGDESRVLTGQVQKHEQDPTTSPAFSPFVSLEAKWMLISLAIFY 480
QY 481 POSTYGSSTGPVVSIVLNVNAPAGAVAPSLWPKVILKCPPLSTLYSKTFEVL 540
DB 481 POSTYGSSTGPVVSIVLNVNAPAGAVAPSLWPKVILKCPPLSTLYSKTFEVL 540
QY 541 LPKRLSEWIACTEAWTLYRSTLAAAGLVNAAQHEVALGLYLSLAAQVGLSAAVAV 600
DB 541 LPKRLSEWIACTEAWTLYRSTLAAAGLVNAAQHEVALGLYLSLAAQVGLSAAVAV 600
QY 601 LAHESALAILFLLIYPARAHITFDYPTFWGLGLQVAFQSTVARELQPLKMKVGTTEL 660
DB 601 LAHESALAILFLLIYPARAHITFDYPTFWGLGLQVAFQSTVARELQPLKMKVGTTEL 660

```

QY 181 NYAGYVPAPATIPYPIPLVNAVAGYATISFWMQQTITPTSVIMNSITSHWFLVQVPT 240
 DB 181 NYAGYVPAPATIPYPIPLVNAVAGYATISFWMQQTITPTSVIMNSITSHWFLVQVPT 240
 QY 241 ASFLVTPSERHYPNQGWPSVETSCVAEEFATSGVLMVLTGSLVNSYNTPTYSALGLL 300
 DB 241 ASFLVTPSERHYPNQGWPSVETSCVAEEFATSGVLMVLTGSLVNSYNTPTYSALGLL 300
 QY 301 DFALFLYBPNNITPGNTNTPVSYSSSTAFHFLPCAGDTAPLTTAATPEMKKLYETSTNG 360
 DB 301 DFALFLYBPNNITPGNTNTPVSYSSSTAFHFLPCAGDTAPLTTAATPEMKKLYETSTNG 360
 QY 361 VETGRCIATLFLNLAFTLGGRLPTLTLSSAGQLFYSPPVVSANSEPIVKLYTSVENAQ 420
 DB 361 VETGRCIATLFLNLAFTLGGRLPTLTLSSAGQLFYSPPVVSANSEPIVKLYTSVENAQ 420
 QY 421 QERGIATPHGLGESHVVLQDYUNQHEQGRPIEPSPAPSKGPSVLKANIWIWLSIAAAY 480
 DB 421 QERGIATPHGLGESHVVLQDYUNQHEQGRPIEPSPAPSKGPSVLKANIWIWLSIAAAY 480
 QY 481 DQSTYGSSTGPPVVSSTLVNATAGAVAPSLIMWFLVLLGKPLSTLQNSKTEPVLP 540
 DB 481 DQSTYGSSTGPPVVSSTLVNATAGAVAPSLIMWFLVLLGKPLSTLQNSKTEPVLP 540
 QY 541 LPKKLSWEAGTTKAGYPYNYNTTASQQLLVENAAHBPVAISTYTLISLWGHVSTSAVAV 600
 DB 541 LPKKLSWEAGTTKAGYPYNYNTTASQQLLVENAAHBPVAISTYTLISLWGHVSTSAVAV 600
 QY 601 LAPHSALALLEDTLYPAPAHTEFQSPPEPLGLGSAFESLVAHQPLKMKVSKTEFL 660
 DB 601 LAPHSALALLEDTLYPAPAHTEFQSPPEPLGLGSAFESLVAHQPLKMKVSKTEFL 660

RESULT 10

AAW76369

ID AAW76369 standard; Protein: 660 AA.

XX AAW76369;

XX 03-DEC-1998 (first entry)

DE Hepatitis E virus hollow particle protein #2.

XX Hollow particle protein, virus, antibody, detection, immunodassay.

FW infection

OS Hepatitis virus.

XX JF10234383-A

XX 08-SEP 1998

XX 28-FEB-1997; 97JP-0062445.

XX 28-FEB-1997; 97JP-0062445

XX (DENK-) DENKA SEIKEN KK.

PA (KORU-) KOKUTITSU YAKO EISEI KENKYUSHO.

XX WPI: 1998-535037/46.

XX N-PSDB: AAW61688.

XX Hepatitis E virus hollow particle polypeptide(s) and nucleic acids
 PT encoding it, useful for more accurate detection of HEV in samples,
 PT using immunoassays and nucleic acid hybridisation

PS Claim 13; Page 24-26; 29pp; Japanese.

XX This sequence represents a Hepatitis E viral hollow particle protein.
 CC This polypeptide can be used to raise antibodies to detect HEV infection
 CC in samples, e.g. by immuno assay based techniques, and the nucleic acid
 CC can be used for the same in nucleic acid hybridisation assays. The
 CC polypeptides and nucleic acids allow more accurate detection of HEV than

CC previously possible.

XX Sequence 560 AA;

Query Match 84.7%; Score 559; DB 19; Length 660;
 Best local Similarity 99.8%; Pred. No. 0;
 Matches 659; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPPEFTLLILMLPLMIPAFITTFQFSPEEPFSSSSSSSSSTWSSAVAGTEFTLTPVIMV 60
 DB 1 MPPEFTLLILMLPLMIPAFITTFQFSPEEPFSSSSSSSSSTWSSAVAGTEFTLTPVIMV 60
 QY 61 PFAPDVTAAAGAGPPVVSQAPPIGSAWPTAGAPPAVAVASPPPTATATAHATAVAAHPTP 120
 DB 61 PFAPDVTAAAGAGPPVVSQAPPIGSAWPTAGAPPAVAVASPPPTATATAHATAVAAHPTP 120
 QY 121 FVLVVSERGALLERLYMLSTLTSVATGTNVLVYAAPLSTLELQNTTHIMAEAS 180
 DB 121 FVLVVSERGALLERLYMLSTLTSVATGTNVLVYAAPLSTLELQNTTHIMAEAS 180
 QY 181 NYAGYVPAPATIPYPIPLVNAVAGYATISFWMQQTITPTSVIMNSITSHWFLVQVPT 240
 DB 181 NYAGYVPAPATIPYPIPLVNAVAGYATISFWMQQTITPTSVIMNSITSHWFLVQVPT 240
 QY 241 ASFLVTPSERHYPNQGWPSVETSCVAEEFATSGVLMVLTGSLVNSYNTPTYSALGLL 300
 DB 241 ASFLVTPSERHYPNQGWPSVETSCVAEEFATSGVLMVLTGSLVNSYNTPTYSALGLL 300
 QY 301 DFALFLYBPNNITPGNTNTPVSYSSSTAFHFLPCAGDTAPLTTAATPEMKKLYETSTNG 360
 DB 301 DFALFLYBPNNITPGNTNTPVSYSSSTAFHFLPCAGDTAPLTTAATPEMKKLYETSTNG 360
 QY 361 VETGRCIATLFLNLAFTLGGRLPTLTLSSAGQLFYSPPVVSANSEPIVKLYTSVENAQ 420
 DB 361 VETGRCIATLFLNLAFTLGGRLPTLTLSSAGQLFYSPPVVSANSEPIVKLYTSVENAQ 420
 QY 421 QERGIATPHGLGESHVVLQDYUNQHEQGRPIEPSPAPSKGPSVLKANIWIWLSIAAAY 480
 DB 421 QERGIATPHGLGESHVVLQDYUNQHEQGRPIEPSPAPSKGPSVLKANIWIWLSIAAAY 480
 QY 481 DQSTYGSSTGPPVVSSTLVNATAGAVAPSLIMWFLVLLGKPLSTLQNSKTEPVLP 540
 DB 481 DQSTYGSSTGPPVVSSTLVNATAGAVAPSLIMWFLVLLGKPLSTLQNSKTEPVLP 540
 QY 541 LFKKLSWEAGTTKAGYPYNYNTTASQQLLVENAAHBPVAISTYTLISLWGHVSTSAVAV 600
 DB 541 LFKKLSWEAGTTKAGYPYNYNTTASQQLLVENAAHBPVAISTYTLISLWGHVSTSAVAV 600
 QY 601 LAPHSALALLEDTLYPAPAHTEFQSPPEPLGLGSAFESLVAHQPLKMKVSKTEFL 660
 DB 601 LAPHSALALLEDTLYPAPAHTEFQSPPEPLGLGSAFESLVAHQPLKMKVSKTEFL 660

RESULT 11

AAW71210

ID AAW71210 standard; Protein: 660 AA.

XX AAW71210;

XX 25-MAR-2003 (updated)

XX 30-OCT-1998 (first entry)

XX Protein encoded by ORF 3 of the Japanese isolate of PT-NANP.

XX Encoded, transmitted, and/or hepatitis virus, blood collection;
 KW HEV; PT-NANP; detection; vaccine.
 XX Hepatitis virus.

XX Key location/qualifiers

XX Misc-difference 496

XX /note- "not specified"

XX

QY 242 PYTGALGLIDPILALHLEHLENTIPENTNTPVSRYSSTAPHPILPBGAGGTAELETTTAAATPEMK 351
DB 181 PYTGALGLIDPILALHLEHLENTIPENTNTPVSRYSSTAPHPILPBGAGGTAELETTTAAATPEMK 240
QY 452 DLYFTSTNGVGFIRGIALTFENIADTILGGIPTELISSAGGOLFYSRPVVSANGEPYVK 411
DB 241 DLYFTSTNGVGFIRGIALTFENIADTILGGIPTELISSAGGOLFYSRPVVSANGEPYVK 300
QY 412 IYTSVENNAQQDPKGIATIPEDITGIESPVVILQYVINGHEPQPTSPDAPSPFESVLPAIVL 471
DB 391 IYTSVENNAQQDPKGIATIPEDITGIESPVVILQYVINGHEPQPTSPDAPSPFESVLPAIVL 360
QY 472 WLSITAAFYDQSTVGSSTGPPVYVSDSVTLVNVATGAGAVAPSIDWTKVTLDCRPILSTIQ 531
DB 461 WLSITAAFYDQSTVGSSTGPPVYVSDSVTLVNVATGAGAVAPSIDWTKVTLDCRPILSTIQ 420
QY 542 YSKTEFFVILPGRKLSFWEAGTTRKAGYPYNYNTTASDQLLVENNAAGHRVAISTYTTSLGAG 591
DB 421 YSKTEFFVILPGRKLSFWEAGTTRKAGYPYNYNTTASDQLLVENNAAGHRVAISTYTTSLGAG 480
QY 592 PVSISAVAVIAPHSAIALEEDTLDYPARAHTFDDEPFPKPLGLQ 636
DB 481 PVSISAVAVIAPHSAIALEEDTLDYPARAHTFDDEPFPKPLGLQ 525

Search completed: August 5, 2003, 09:32:49
Job time : 49 secs

GenCode version 5.1.6
Copyright (c) 1994-2003 Compugen Ltd.
OM protein protein search, using sw model
Run on: August 5, 2003, 09:41:39 ; Search time 19 seconds
(without alignments)
1469,746 Million cell updates/sec
Filter: US 09 851 410A H
Percent score: 66.0
Sequences: 1 MKP6HLLIMLPMFAP.....QSTVARLQLRMKVKTRFL 666
Scoring table: 0.1176
Gapop 60.0 ; Gapext 60.0
Scor test: 428/17 seqs, 4240858 residues
Word size: 1

Total number of hits satisfying chosen parameters: 428717

Minimum hit seq length: 0

Maximum hit seq length: 260909966

Post processing: listing first 45 summaries

Database: Issued Patents, AA.*
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 25


```

1  FILING DATE: 17-JUN-1988
2  ATTORNEY/AGENT INFORMATION:
3  NAME: Scott, Charles K.
4  REGISTRATION NUMBER: 48,615
5  REFERENCE/EXHIBIT NUMBER: 4,600-0184,24
6  TELECOMMUNICATION INFORMATION:
7  TELEPHONE: (650) 424-0880
8  TELEPHONE: (650) 424-0960
9  TELEFAX: (650) 424-0960
10 INFORMATION FOR SEQ ID NO: 4:
11 LENGTH: 663 amino acids
12 TYPE: amino acid
13 FEATURE: linear
14 MOLECULE TYPE: protein
15 US 09 424 644 14

```

```

Query Match 100.0% 663 663 0 0 163 163 663
Best Local Similarity 100.0% Ident. No. 0
Matches 660 Conservative 0 Mismatches 0 Indels 0 Gaps 0

```

```

Q7 1 MEKPRPRLLELMFIMLAPPPQSSRRPRSSSSGPPWGLRVLSQFAIYDPTIN 60
16 1 MEKPRPRLLELMFIMLAPPPQSSRRPRSSSSGPPWGLRVLSQFAIYDPTIN 60
Q7 61 PEAPGVTAAGAAGPPVGPAPPLASAMPQAGRAVASRPPPTATGAAPLAVAPADTP 120
16 61 PEAPGVTAAGAAGPPVGPAPPLASAMPQAGRAVASRPPPTATGAAPLAVAPADTP 120
Q7 121 PWDWGRALLPFGVRIQGGGTSNVAVATINI VIYAASPLPLFPGTNTIHMAEAS 180
16 121 PWDWGRALLPFGVRIQGGGTSNVAVATINI VIYAASPLPLFPGTNTIHMAEAS 180
Q7 181 NYAGVAVATIRYRPIVGNAGGYAISTFWPTTTPTSVOMNSITSTQVRLVQPGI 240
16 181 NYAGVAVATIRYRPIVGNAGGYAISTFWPTTTPTSVOMNSITSTQVRLVQPGI 240
Q7 241 ASELVPSPELHYNQWPSVETSCVAPFFATSCGLVMPCHGSSIVNSYTNIPYTGALL 300
16 241 ASELVPSPELHYNQWPSVETSCVAPFFATSCGLVMPCHGSSIVNSYTNIPYTGALL 300
Q7 601 DFALLEPRLNLPNTINRVSKYSTAHRIKRCAGCTAFITTAATFPKRLDYFTSTNG 360
16 601 DFALLEPRLNLPNTINRVSKYSTAHRIKRCAGCTAFITTAATFPKRLDYFTSTNG 360
Q7 361 VKEELKPEALTEALTEALTEALTEALTEALTEALTEALTEALTEALTEALTEAL 420
16 361 VKEELKPEALTEALTEALTEALTEALTEALTEALTEALTEALTEALTEALTEAL 420
Q7 421 LKRLHATBRLDEPSVVLGVNQNQPPPPSPAPSPSPSPSPSPSPSPSPSPSPSPSP 480
16 421 LKRLHATBRLDEPSVVLGVNQNQPPPPSPAPSPSPSPSPSPSPSPSPSPSPSPSP 480
Q7 481 QSTVSSSHGVVSVSSVHNNVAAGAAVAPSTLQWLYLQSPSTLQSSKFLVLP 540
16 481 QSTVSSSHGVVSVSSVHNNVAAGAAVAPSTLQWLYLQSPSTLQSSKFLVLP 540
Q7 541 LPEKLTWEATTEAWTPRYRNTLAEQGLVERAAERVALSYTHLEQADNLSAVAY 600
16 541 LPEKLTWEATTEAWTPRYRNTLAEQGLVERAAERVALSYTHLEQADNLSAVAY 600
Q7 601 LAPESALALLEHLLYPARAHLEDDPQEPHPLGLQCTAPQSTVAERQPLKRVVTPPI 660
16 601 LAPESALALLEHLLYPARAHLEDDPQEPHPLGLQCTAPQSTVAERQPLKRVVTPPI 660

```

```

RESULT 4
US 09 424 644 14
Sequence 14 Application US/09542644
Patent No. 6,214,928
GENERAL INFORMATION
APPLICANT: Scott, Charles K.
ATTORNEY: Barbood, Patricia

```

```

1  APPLICANT: Zhang, Yilan
2  TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENIC AND OTHER THEREOF
3  NUMBER OF SEQUENCES: 1
4  CORRESPONDENT ADDRESS:
5  ADDRESS: Dadi Import & Associates
6  STREET: 350 Cambridge Ave., Suite 250
7  CITY: Palo Alto
8  STATE: CA
9  COUNTRY: USA
10 ZIP: 94306
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: floppy disk
13 COMPUTER: IBM PC compatible
14 OPERATING SYSTEM: PC DOS/MS DOS
15 SOFTWARE: Patent In Release #100, Version #1.25
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: 09/0942644
18 FILING DATE:
19 CLASSIFICATION: 4.0
20 ATTORNEY/AGENT INFORMATION:
21 NAME: Fadiat, Gary R.
22 REGISTRATION NUMBER: 48,615
23 REFERENCE/EXHIBIT NUMBER: 4,600-0294,40
24 TELECOMMUNICATION INFORMATION:
25 TELEPHONE: (415) 424-0880
26 TELEFAX: (415) 424-0960
27 INFORMATION FOR SEQ ID NO: 1:
28 SEQUENCE CHARACTERISTICS:
29 LENGTH: 663 amino acids
30 TYPE: amino acid
31 TOPOLOGY: linear
32 MOLECULE TYPE: protein
33 HYPOTHETICAL: N
34 ORIGINAL SOURCE:
35 INDIVIDUAL ISOLATE: Hepatitis E Virus (Barbood)
36 INDIVIDUAL ISOLATE: REF 2
37 US 09 542 634 14

```

```

Query Match 100.0% 663 663 0 0 163 163 663
Best Local Similarity 100.0% Ident. No. 0
Matches 660 Conservative 0 Mismatches 0 Indels 0 Gaps 0
Q7 1 MEKPRPRLLELMFIMLAPPPQSSRRPRSSSSGPPWGLRVLSQFAIYDPTIN 60
16 1 MEKPRPRLLELMFIMLAPPPQSSRRPRSSSSGPPWGLRVLSQFAIYDPTIN 60
Q7 61 PEAPGVTAAGAAGPPVGPAPPLASAMPQAGRAVASRPPPTATGAAPLAVAPADTP 120
16 61 PEAPGVTAAGAAGPPVGPAPPLASAMPQAGRAVASRPPPTATGAAPLAVAPADTP 120
Q7 121 PWDWGRALLPFGVRIQGGGTSNVAVATINI VIYAASPLPLFPGTNTIHMAEAS 180
16 121 PWDWGRALLPFGVRIQGGGTSNVAVATINI VIYAASPLPLFPGTNTIHMAEAS 180
Q7 181 NYAGVAVATIRYRPIVGNAGGYAISTFWPTTTPTSVOMNSITSTQVRLVQPGI 240
16 181 NYAGVAVATIRYRPIVGNAGGYAISTFWPTTTPTSVOMNSITSTQVRLVQPGI 240
Q7 241 ASELVPSPELHYNQWPSVETSCVAPFFATSCGLVMPCHGSSIVNSYTNIPYTGALL 300
16 241 ASELVPSPELHYNQWPSVETSCVAPFFATSCGLVMPCHGSSIVNSYTNIPYTGALL 300
Q7 601 DFALLEPRLNLPNTINRVSKYSTAHRIKRCAGCTAFITTAATFPKRLDYFTSTNG 360
16 601 DFALLEPRLNLPNTINRVSKYSTAHRIKRCAGCTAFITTAATFPKRLDYFTSTNG 360
Q7 361 VKEELKPEALTEALTEALTEALTEALTEALTEALTEALTEALTEALTEALTEAL 420
16 361 VKEELKPEALTEALTEALTEALTEALTEALTEALTEALTEALTEALTEALTEAL 420
Q7 421 LKRLHATBRLDEPSVVLGVNQNQPPPPSPAPSPSPSPSPSPSPSPSPSPSPSPSP 480
16 421 LKRLHATBRLDEPSVVLGVNQNQPPPPSPAPSPSPSPSPSPSPSPSPSPSPSPSP 480
Q7 481 QSTVSSSHGVVSVSSVHNNVAAGAAVAPSTLQWLYLQSPSTLQSSKFLVLP 540
16 481 QSTVSSSHGVVSVSSVHNNVAAGAAVAPSTLQWLYLQSPSTLQSSKFLVLP 540
Q7 541 LPEKLTWEATTEAWTPRYRNTLAEQGLVERAAERVALSYTHLEQADNLSAVAY 600
16 541 LPEKLTWEATTEAWTPRYRNTLAEQGLVERAAERVALSYTHLEQADNLSAVAY 600
Q7 601 LAPESALALLEHLLYPARAHLEDDPQEPHPLGLQCTAPQSTVAERQPLKRVVTPPI 660
16 601 LAPESALALLEHLLYPARAHLEDDPQEPHPLGLQCTAPQSTVAERQPLKRVVTPPI 660

```

QY 441 DQSYGSGTGPVVYSDVTLVNVATCAQAVAPSIQWTKVTLGCPPLSTIQYSGTFFVLP 540
DB 441 DQSYGSGTGPVVYSDVTLVNVATCAQAVAPSIQWTKVTLGCPPLSTIQYSGTFFVLP 540
QY 541 LRGLSPWEAGTITKAGYPYNYNTASQQLLVNAAGHRVALISFYTISLAGPVSISAVAV 600
DB 541 LRGLSPWEAGTITKAGYPYNYNTASQQLLVNAAGHRVALISFYTISLAGPVSISAVAV 600
QY 601 LAPHSALALTDITLDPAPAHATFDGCPPEPFGHGGCAFQSTVAFIOLPKMFVSKTRPL 660
DB 601 LAPHSALALTDITLDPAPAHATFDGCPPEPFGHGGCAFQSTVAFIOLPKMFVSKTRPL 660

RESULT 7

US-09-128-275A-8
Sequence 8, Application US/0910827A
Patent No. 629005
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R
APPLICANT: Yarbough, Patrice O
APPLICANT: Bradley, Daniel W
APPLICANT: Friedman, Kirk
APPLICANT: Tam, Albert
APPLICANT: Fry, Kirk E
TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-WINDOWS-NT
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/128,275A
FILING DATE: 03 AUG 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08,079,823
FILING DATE: 25-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07,681,078
FILING DATE: 05-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07,505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07,420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07,467,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07,336,672
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07,208,997
FILING DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Petithory, Joanne R.
REGISTRATION NUMBER: 42,995
REFERENCE/PATENT NUMBER: 4600-0183.24
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:

LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-128-275A-8

Query Match: 100.0% Score 660; DB 8; Length 660;
Exact Local Similarity: 100.0%; Prod. No. 0;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLMLPLMLPFAIPFPQPSGRGRGRSGSGSGPWRKVRKVSQPAIPYTHPTN 60
DB 1 MRPRPILLLMLPLMLPFAIPFPQPSGRGRGRSGSGSGPWRKVRKVSQPAIPYTHPTN 60
QY 61 PFAPVTAAAGATGPRVQGFARHTGASAKHRCAGRPAAVABRRPILATAGAPLAAVAHAHP 120
DB 61 PFAPVTAAAGATGPRVQGFARHTGASAKHRCAGRPAAVABRRPILATAGAPLAAVAHAHP 120
QY 121 PVPATGEGATGAKRCYRCESTGRTTSVATGRTTGVATGRTTGVATGRTTGVATGRTTGVATG 180
DB 121 PVPATGEGATGAKRCYRCESTGRTTSSVAUGTNILVLYAAHPLPLPLQKTRTHHMAILAS 180
QY 141 NTAATVAGATIRYRTLVENAVATVATGRTTGVATGRTTGVATGRTTGVATGRTTGVATG 240
DB 141 NTAATVAGATIRYRTLVENAVATVATGRTTGVATGRTTGVATGRTTGVATGRTTGVATG 240
QY 241 ASPIVPSPIPRPGRGKQK 400
DB 241 ASPIVPSPIPRPGRGKQK 400
QY 401 DEAFITFEFRTTNTNTIRVSVSYSTAPHPITPRGAGTAAELTTTAAITFMKILYPTSTG 460
DB 401 DEAFITFEFRTTNTNTIRVSVSYSTAPHPITPRGAGTAAELTTTAAITFMKILYPTSTG 460
QY 461 VEHGEGTALTFENAKHLLGSLPTELSSAGDGLPYKGVYSANPILVELYISVENAO 420
DB 461 VEHGEGTALTFENAKHLLGSLPTELSSAGDGLPYKGVYSANPILVELYISVENAO 420
QY 421 QMGKATPRHGRHGRSVVAVQV 480
DB 421 QMGKATPRHGRHGRSVVAVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 480
QY 481 DQSYGSGTGPVVYSDVTLVNVATCAQAVAPSIQWTKVTLGCPPLSTIQYSGTFFVLP 540
DB 481 DQSYGSGTGPVVYSDVTLVNVATCAQAVAPSIQWTKVTLGCPPLSTIQYSGTFFVLP 540
QY 541 LRGLSPWEAGTITKAGYPYNYNTASQQLLVNAAGHRVALISFYTISLAGPVSISAVAV 600
DB 541 LRGLSPWEAGTITKAGYPYNYNTASQQLLVNAAGHRVALISFYTISLAGPVSISAVAV 600
QY 601 LAPHSALALTDITLDPAPAHATFDGCPPEPFGHGGCAFQSTVAFIOLPKMFVSKTRPL 660
DB 601 LAPHSALALTDITLDPAPAHATFDGCPPEPFGHGGCAFQSTVAFIOLPKMFVSKTRPL 660

RESULT 8

US-09-477-292-14
Sequence 14, Application US/09177292
Patent No. 6291641
GENERAL INFORMATION:
APPLICANT: Fuotsl, Thomas
APPLICANT: McAtee, Patrick
APPLICANT: Yarbough, Patrice
APPLICANT: Zhang, Yitao
TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Catleco M. Desjardins, M.D.
STREET: 505 Petroscoot Drive
CITY: Redwood City
STATE: CA
COUNTRY: US
ZIP: 94065

```

1 COMPUTER READABLE FORM:
2 MEDIUM TYPE: floppy disk
3
4 OPERATING SYSTEM: PC-DOS/MS DOS
5
6 SOFTWARE: Patent to Release #1.0, Version #1.25
7
8 CURRENT APPLICATION DATA:
9
10 APPLICATION NUMBER: US/08/477,292
11 FILING DATE:
12
13 CLASSIFICATION:
14
15 PRIOR APPLICATION DATA:
16
17 APPLICATION NUMBER: US/08/427,427
18 FILING DATE:
19
20 ATTORNEY/AGENT INFORMATION:
21
22 NAME: Brookes, Allan A.
23 REGISTRATION NUMBER: 46,474
24 REFERENCE/POCKET NUMBER: 6495
25
26 TELECOMMUNICATION INFORMATION:
27
28 TELEPHONE: (415) 369-9500
29 TELEFAX: (415) 668-0709
30
31 INFORMATION FOR SEQ ID NO: 1:
32
33 SEQUENCE CHARACTERISTICS:
34
35 LENGTH: 660 amino acids
36 TYPE: amino acids
37 TOPOLOGY: linear
38
39 INDIVIDUAL ISOLATE: Hepatitis E Virus (Huma Strain)
40
41 INDIVIDUAL ISOLATE: ORF-2
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2615
2616
2617
2618
2619
2620
2621
2622

```



```

SEQUENCE SYSTEM: 10/10/MS/MS/MS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US 505,898
  FILING DATE: 20 APRIL 1992
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 422,445
  FILING DATE: 17 JAN 1992
  PRIORITY CLAIM: 3
APPLICATION NUMBER: US 505,898
FILING DATE: 05 APRIL 1999
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 422,445
  FILING DATE: 15 OCTOBER 1989
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 467,486
  FILING DATE: 16 JUNE 1989
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 439,672
  FILING DATE: 11 APRIL 1989
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 208,997
  FILING DATE: 17 JUNE 1988
ATTORNEY/AGENT INFORMATION:
  NAME: Fabiano, Gary R.
  REGISTRATION NUMBER: 44,875
  REFERENCE/BOOKLET NUMBER: 4,600 0/994.40
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (415) 424-0980
  TELEFAX: (415) 424-0960
  INFORMATION FOR SEQ ID NO: 19:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 660 amino acids
  TYPE: amino acid
  TOPOLOGY: Linear
  MOLECULE TYPE: protein
  HYDROPHILIC: No
  ORIGINAL SOURCE:
  INDIVIDUAL ISOLATE: OFF 2, BORMA, FIGURE 9
US 07 851-410A 19

Query Match: 100.00% Score: 660; DB 4; Length: 660;
Best Local Similarity: 100.00% Prod. No. 0;
Matches: 660; Conservative: 0; Mismatches: 0; Gaps: 0;

Q7 1 MRPPPTLLLLMLPMLPAHPGQPSGRGGRGSGSGGFWCDRWISQPPALPYTHPTN 60
D6 1 MRPPPTLLLLMLPMLPAHPGQPSGRGGRGSGSGGFWCDRWISQPPALPYTHPTN 60
Q7 61 PFAPAVTAAGAGIRKVPQAPRPGASAPRPGQAPRPAVASRRPTTAGAAPLTAVAPAHPTP 120
D6 61 PFAPAVTAAGAGIRKVPQAPRPGASAPRPGQAPRPAVASRRPTTAGAAPLTAVAPAHPTP 120
Q7 121 FVPAVASGALLRRPNLSISLPSVAIGTNLVYAAPLSPILPLDQGNTHIMATKAS 180
D6 121 FVPAVASGALLRRPNLSISLPSVAIGTNLVYAAPLSPILPLDQGNTHIMATKAS 180
Q7 181 NYALRYVARALIPRYPLPAPNAVAYALISLSPFQHTTPTSLVMSNSTSDVRLVQPGI 240
D6 181 NYALRYVARALIPRYPLPAPNAVAYALISLSPFQHTTPTSLVMSNSTSDVRLVQPGI 240
Q7 241 ASSELVTSERPHYNGQWRSVETSWAEDEATSGVLMGIDGSLVNSYNTPTTGALAIL 300
D6 241 ASSELVTSERPHYNGQWRSVETSWAEDEATSGVLMGIDGSLVNSYNTPTTGALAIL 300
Q7 301 LFALELEPNLIPNTNIPVRSYSSIAPIHRLRGACGTAGLTITTAATREPMKLYETSTNG 360
D6 301 LFALELEPNLIPNTNIPVRSYSSIAPIHRLRGACGTAGLTITTAATREPMKLYETSTNG 360
Q7 361 VGEIIPGALITLNLALILDLGLPTELSSAGQLPYSPVSSANGDEPTVKLYTSVENAQ 420
D6 361 VGEIIPGALITLNLALILDLGLPTELSSAGQLPYSPVSSANGDEPTVKLYTSVENAQ 420

SEQUENCE 14, Application: 10/10/MS/MS/MS
GENERAL INFORMATION:
  APPLICANT:
  TITLE OF INVENTION: HETEROLOGOUS VIBRIO ANTIGENS AND
  TITLE OF INVENTION: USES THEREOF
  NUMBER OF SEQUENCES: 41
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Dechinger & Associates
  STREET: P.O. Box 60000
  CITY: Palo Alto
  STATE: CA
  COUNTRY: USA
  ZIP: 94306
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patent in Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: 10/10/MS/MS/MS
  FILING DATE:
  CLASSIFICATION:
  ATTORNEY/AGENT INFORMATION:
  NAME: Fabiano, Gary R.
  REGISTRATION NUMBER: 44,875
  REFERENCE/BOOKLET NUMBER: 4,600 0/994.41
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (415) 424-0980
  TELEFAX: (415) 424-0960
  INFORMATION FOR SEQ ID NO: 19:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 660 amino acids
  TYPE: amino acid
  TOPOLOGY: Linear
  MOLECULE TYPE: protein
  HYDROPHILIC: No
  ORIGINAL SOURCE:
  INDIVIDUAL ISOLATE: OFF 2
  INDIVIDUAL ISOLATE: Bormann, E. Virus (Patent in Release)
  US 09-851-410A-14

Query Match: 100.00% Score: 660; DB 4; Length: 660;
Best Local Similarity: 100.00% Prod. No. 0;
Matches: 660; Conservative: 0; Mismatches: 0; Gaps: 0;

Q7 1 MRPPPTLLLLMLPMLPAHPGQPSGRGGRGSGSGGFWCDRWISQPPALPYTHPTN 60
D6 1 MRPPPTLLLLMLPMLPAHPGQPSGRGGRGSGSGGFWCDRWISQPPALPYTHPTN 60
Q7 61 PFAPAVTAAGAGIRKVPQAPRPGASAPRPGQAPRPAVASRRPTTAGAAPLTAVAPAHPTP 120
D6 61 PFAPAVTAAGAGIRKVPQAPRPGASAPRPGQAPRPAVASRRPTTAGAAPLTAVAPAHPTP 120

```


1 ZIPS: 4000
 2 COMPUTER READABLE FORM:
 3 MEDIUM TYPE: floppy disk
 4 COMPUTER: IBM PC compat 160
 5 OPERATING SYSTEM: PC-DOS/MS-DOS
 6 SOFTWARE: Patient In Release #1.0, Version #1.25
 7 CURRENT APPLICATION DATA:
 8 APPLICATION NUMBER: 05/08/477, 492
 9 FILING DATE:
 10 CLASSIFICATION:
 11 PRIOR APPLICATION DATA:
 12 APPLICATION NUMBER: 05/08/427, 952
 13 FILING DATE:
 14 ATTORNEY/AGENT INFORMATION:
 15 NAME: Brooks, Allan A.
 16 REGISTRATION NUMBER: 46, 674
 17 TELEPHONE: (415) 369, 5215
 18 TELECOMMUNICATION INFORMATION:
 19 TELEPHONE: (415) 369, 9500
 20 TELEFAX: (415) 368, 0709
 21 INFORMATION FOR SEQ ID NO: 15:
 22 SEQUENCE CHARACTERISTICS:
 23 LENGTH: 549 amino acids
 24 TYPE: amino acid
 25 STRANDEDNESS: single
 26 TOPOLOGY: linear
 27 MOLECULE TYPE: protein
 28 HYDROPHILIC: No
 29 ORIGINAL SOURCE:
 30 INDIVIDUAL ISOLATE: Hepatitis E Virus (Basma strain)
 31 INDIVIDUAL ISOLATE: F62kda, FIGURE 4
 32 SEQ ID 477-292 15

Query Match: 84.2% Score: 549; 146 4; Length: 549;

Best Local Similarity: 100.0%; Ident. No. 0;
Matches: 549; Conservation: 0; Mismatches: 0; Indels: 0; Gaps: 0;

112 AVAFAHDPFVFWDSRCALLRQYNLSTSLSSVATGILVLYAAPSPLPLPQGTIN 171
 1 AVAFAHDPFVFWDSRCALLRQYNLSTSLSSVATGILVLYAAPSPLPLPQGTIN 40
 172 THMAHAGVAVAVAKATIRYPLVNAVGVYALSTSWPQFTTPTSVIMNSTISD 241
 1 THMAHAGVAVAVAKATIRYPLVNAVGVYALSTSWPQFTTPTSVIMNSTISD 241
 61 THMAHAGVAVAVAKATIRYPLVNAVGVYALSTSWPQFTTPTSVIMNSTISD 120
 272 VLLVAGTACATVTPRERHRCQKSWETGVAHREATSLVYVLCRGGSLVLYCT 251
 1 VLLVAGTACATVTPRERHRCQKSWETGVAHREATSLVYVLCRGGSLVLYCT 180
 274 FYTGADGILDFALRFRNLPGNTNTRVSPYSTAHPLRRCAGDGTAEITTAATRMK 451
 1 FYTGADGILDFALRFRNLPGNTNTRVSPYSTAHPLRRCAGDGTAEITTAATRMK 240
 372 DLYETSLRWELRGLALLRPLADILLGLLTELSSAGGGLFYSRWVVSANGRPV 411
 241 DLYETSLRWELRGLALLRPLADILLGLLTELSSAGGGLFYSRWVVSANGRPV 400
 472 LYSVENAGQPTLADPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSP 471
 471 LYSVENAGQPTLADPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSP 460
 472 WSLIAARYDQSTKQSTGQVYVGVSTLVNVAALAVAGVAVAVAVAVAVAV 541
 461 WSLIAARYDQSTKQSTGQVYVGVSTLVNVAALAVAGVAVAVAVAVAVAV 420
 542 YSTGFRVLDPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSPSPSP 541
 461 YSTGFRVLDPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSPSPSP 460
 542 YSTGFRVLDPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSPSPSP 541
 461 YSTGFRVLDPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSPSPSP 460
 542 YSTGFRVLDPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSPSPSP 541
 461 YSTGFRVLDPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSPSPSP 460

05 MKVVKRFFL 660
 1111111111
 041 MKVVKRFFL 549
 1111111111
 112 AVAFAHDPFVFWDSRCALLRQYNLSTSLSSVATGILVLYAAPSPLPLPQGTIN 171
 1 AVAFAHDPFVFWDSRCALLRQYNLSTSLSSVATGILVLYAAPSPLPLPQGTIN 40
 172 THMAHAGVAVAVAKATIRYPLVNAVGVYALSTSWPQFTTPTSVIMNSTISD 241
 1 THMAHAGVAVAVAKATIRYPLVNAVGVYALSTSWPQFTTPTSVIMNSTISD 241
 61 THMAHAGVAVAVAKATIRYPLVNAVGVYALSTSWPQFTTPTSVIMNSTISD 120
 272 VLLVAGTACATVTPRERHRCQKSWETGVAHREATSLVYVLCRGGSLVLYCT 251
 1 VLLVAGTACATVTPRERHRCQKSWETGVAHREATSLVYVLCRGGSLVLYCT 180
 274 FYTGADGILDFALRFRNLPGNTNTRVSPYSTAHPLRRCAGDGTAEITTAATRMK 451
 1 FYTGADGILDFALRFRNLPGNTNTRVSPYSTAHPLRRCAGDGTAEITTAATRMK 240
 372 DLYETSLRWELRGLALLRPLADILLGLLTELSSAGGGLFYSRWVVSANGRPV 411
 241 DLYETSLRWELRGLALLRPLADILLGLLTELSSAGGGLFYSRWVVSANGRPV 400
 472 LYSVENAGQPTLADPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSP 471
 471 LYSVENAGQPTLADPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSP 460
 472 WSLIAARYDQSTKQSTGQVYVGVSTLVNVAALAVAGVAVAVAVAVAVAV 541
 461 WSLIAARYDQSTKQSTGQVYVGVSTLVNVAALAVAGVAVAVAVAVAVAV 420
 542 YSTGFRVLDPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSPSPSP 541
 461 YSTGFRVLDPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSPSPSP 460
 542 YSTGFRVLDPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSPSPSP 541
 461 YSTGFRVLDPRHLCSPSVVLDGVNQHGLPPTSPSPSPSPSPSPSPSPSP 460

Query Match: 84.2% Score: 549; 146 4; Length: 549;
 Best Local Similarity: 100.0%; Ident. No. 0;
 Matches: 549; Conservation: 0; Mismatches: 0; Indels: 0; Gaps: 0;

```

QY 412 LYSVFNAAQOKRTAIPHDIDLAGSPVVIQVYDNQHEQDPPTSPAPSPPSVIPANDVI 471
ib 412 LYSVFNAAQOKRTAIPHDIDLAGSPVVIQVYDNQHEQDPPTSPAPSPPSVIPANDVI 471
QY 431 LYSVFNAAQOKRTAIPHDIDLAGSPVVIQVYDNQHEQDPPTSPAPSPPSVLEPANDVL 360
ib 431 LYSVFNAAQOKRTAIPHDIDLAGSPVVIQVYDNQHEQDPPTSPAPSPPSVLEPANDVL 360
QY 472 WLSLIAAEYDQSTYGSSTGSPVYVSDSVTIIVNVAITGAQAVARSIDWTKVTLGGRPLSTIQ 531
ib 472 WLSLIAAEYDQSTYGSSTGSPVYVSDSVTIIVNVAITGAQAVARSIDWTKVTLGGRPLSTIQ 531
QY 461 WLSLIAAEYDQSTYGSSTGSPVYVSDSVTIIVNVAITGAQAVARSIDWTKVTLGGRPLSTIQ 420
ib 461 WLSLIAAEYDQSTYGSSTGSPVYVSDSVTIIVNVAITGAQAVARSIDWTKVTLGGRPLSTIQ 420
QY 542 YSKTFFVLPGLKLSWEAGTTKAGYPYNYNTTASDQLLVENAAGHRVAISTYTTSLGAG 591
ib 542 YSKTFFVLPGLKLSWEAGTTKAGYPYNYNTTASDQLLVENAAGHRVAISTYTTSLGAG 591
QY 421 YSKTFFVLPGLKLSWEAGTTKAGYPYNYNTTASDQLLVENAAGHRVAISTYTTSLGAG 480
ib 421 YSKTFFVLPGLKLSWEAGTTKAGYPYNYNTTASDQLLVENAAGHRVAISTYTTSLGAG 480
QY 592 PVSISAVAVIAPHSAIALLLEDITLDYPARAHTFDDHCPGPIGLGQCAFQSTVAELQRIK 651
ib 592 PVSISAVAVIAPHSAIALLLEDITLDYPARAHTFDDHCPGPIGLGQCAFQSTVAELQRIK 651
QY 481 PVSISAVAVIAPHSAIALLLEDITLDYPARAHTFDDHCPGPIGLGQCAFQSTVAELQRIK 540
ib 481 PVSISAVAVIAPHSAIALLLEDITLDYPARAHTFDDHCPGPIGLGQCAFQSTVAELQRIK 540
QY 652 MKVGKTREL 660
ib 652 MKVGKTREL 660
QY 541 MKVGKTREL 549
ib 541 MKVGKTREL 549

```

Search completed: August 5, 2003, 09:34:36
Job time : 20 secs


```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 660 amino acids
: TYPE: amino acid
: STRANDEDNESS: Hepatitis E Virus (Burma strain)
: ORF-2
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-769-066-13

Query Match          100.0%; Score 660; DB 10; Length 660;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMFIPMI PAPPDQPSQRRPQPSQSGQPCWCPVDSQPPFAIYIDPTN 60
DB 1 MRPRPILLLLMFIPML PAPPDQPSQRRPQPSQSGQPCWCPVDSQPPFAIYIDPTN 60
QY 61 PRAPDVTAAACAGPRVPQAPRI GSAMRQACQAPPAVASPPRPPTTACAAPITAVAFAPHTP 120
DB 61 PRAPDVTAAACAGPRVQAPRI GSAMRQACQAPPAVASPPRPPTTACAAPITAVAFAPHTP 120
QY 121 PVTVISPGAILRPPQYNLSPTSTSSVATGNTNIVLYAAPISPIPIAPGCTNTHIMATFAS 180
DB 121 PVTVDSGAILRPPQYNLSPTSTSSVATGNTNIVLYAAPISPLLPLOGSTNTHIMATFAS 180
QY 181 NYAQYVPAPATDIPYPI VNAVQGYAISTSEWQVTTTTTSVNRNISTETQVRIIVQPGI 240
DB 181 NYAQYVPAPATIPYPI VNAVQGYAISTSEWQVTTTTTSVNRNISTETQVRIIVQPGI 240
QY 241 ASSEVAVISERLHYRNGWKEVETSGVAREEATSLVMKCTHGSIVNVTNITYGAGLGL 300
DB 241 ASHVIIVSEPHVNPNGWSEVETSGVAREEATSLVMKCTHGSIVNSVTNITYGAGLGL 300
QY 301 DEALELERNLTPTNTNTVSVSYSTAPRPLPPLPAPAGTAETITTAATPEMKDLYPTSTNG 360
DB 301 DEALELERNLTPTNTNTVSVSYSTAPRPLPPLPAPAGTAETITTAATPEMKDLYPTSTNG 360
QY 361 VGRIGRIATITPNIADTLIGGLPTEL ISSAGGLFSPVVSANCPTTVKLYTSVENAQ 420
DB 361 VGEIGRIATITPNIADTLIGGLPTEL ISSAGGLFSPVVSANGEPTTVKLYTSVENAQ 420
QY 421 QDKGIAIPIPIIDGESRVI IQVYDNGHQFQRPPTSPAPSPSPSVLERANDVLWLSITAAFY 480
DB 421 QDKGIAIPIPIIDGESRVI IQVYDNGHQFQRPPTSPAPSPSPSVLERANDVLWLSITAAFY 480
QY 481 PQSTYASSTGPPVVSNSVTLVNVATCAQAVASPLDWTVTYDGPPLSTIQGYSTFEVLP 540
DB 481 EQSTYASSTGPPVVSNSVTLVNVATCAQAVASPLDWTVTYDGPPLSTIQGYSTFEVLP 540
QY 541 LRGKLSFWEAGTICAGYPNYNTASDGLLVENAACTHEVAISTYITSLICAPVSIKAVAV 600
DB 541 LRGKLSFWEAGTICAGYPNYNTASDGLLVENAACTHEVAISTYITSLICAPVSIKAVAV 600
QY 601 LARHSALALGDLILYAPAKARTFDGPPPPPIIAGCAPSPSVAPAGPGRKMKVSKTPRI 660
DB 601 LARHSALALGDLILYAPAKARTFDGPPPPPIIAGCAPSPSVAPAGPGRKMKVSKTPRI 660

RESULT 2
US-09-851-410-9
: Sequence 8; Application US/09851410
: Publication No. US20030124510A1
: GENERAL INFORMATION:
: APPLICANT: Reyes, Gregory E
: Yarbough, Patricia O
: Bradley, Daniel W
: Krawczynski, Krzysztof J
: Tam, Albert
: Fry, Kirk E
: TITLE OF INVENTION: DNA Sequences of Enterically Transmitted

```

```

: REFERENCE INFORMATION:
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Techline & Associates
: STREET: 350 Cambridge Avenue, Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US 07/7851410
: FILING DATE: 07-May-2001
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/128,375
: FILING DATE: 04/09/2001
: APPLICATION NUMBER: 07/691,078
: FILING DATE: 05-APR-1991
: APPLICATION NUMBER: US 07/505,888
: FILING DATE: 05-APR-1990
: APPLICATION NUMBER: US 07/420,421
: FILING DATE: 13-OCT-1989
: APPLICATION NUMBER: US 07/667,486
: FILING DATE: 16-JUN-1989
: APPLICATION NUMBER: US 07/779,672
: FILING DATE: 11-APR-1989
: APPLICATION NUMBER: 02/0739,097
: FILING DATE: 17-JUN-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Pettibury, Joanne K.
: PESTICIPATION NUMBER: 42,995
: REFERENCE/DOCKET NUMBER: 4600-0184,24
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 324-0880
: TELEFAX: (650) 324-0960
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 660 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-851 410 8

```

```

Query Match          100.0%; Score 660; DB 11; Length 660;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMFIPMI PAPPDQPSQRRPQPSQSGQPCWCPVDSQPPFAIYIDPTN 60
DB 1 MRPRPILLLLMFIPML PAPPDQPSQRRPQPSQSGQPCWCPVDSQPPFAIYIDPTN 60
QY 61 PRAPDVTAAACAGPRVPQAPRI GSAMRQACQAPPAVASPPRPPTTACAAPITAVAFAPHTP 120
DB 61 PRAPDVTAAACAGPRVQAPRI GSAMRQACQAPPAVASPPRPPTTACAAPITAVAFAPHTP 120
QY 121 PVTVISPGAILRPPQYNLSPTSTSSVATGNTNIVLYAAPISPIPIAPGCTNTHIMATFAS 180
DB 121 PVTVDSGAILRPPQYNLSPTSTSSVATGNTNIVLYAAPISPLLPLOGSTNTHIMATFAS 180
QY 181 NYAQYVPAPATDIPYPI VNAVQGYAISTSEWQVTTTTTSVNRNISTETQVRIIVQPGI 240
DB 181 NYAQYVPAPATIPYPI VNAVQGYAISTSEWQVTTTTTSVNRNISTETQVRIIVQPGI 240
QY 241 ASSEVAVISERLHYRNGWKEVETSGVAREEATSLVMKCTHGSIVNVTNITYGAGLGL 300
DB 241 ASHVIIVSEPHVNPNGWSEVETSGVAREEATSLVMKCTHGSIVNSVTNITYGAGLGL 300
QY 301 DEALELERNLTPTNTNTVSVSYSTAPRPLPPLPAPAGTAETITTAATPEMKDLYPTSTNG 360
DB 301 DEALELERNLTPTNTNTVSVSYSTAPRPLPPLPAPAGTAETITTAATPEMKDLYPTSTNG 360
QY 361 VGRIGRIATITPNIADTLIGGLPTEL ISSAGGLFSPVVSANCPTTVKLYTSVENAQ 420
DB 361 VGEIGRIATITPNIADTLIGGLPTEL ISSAGGLFSPVVSANGEPTTVKLYTSVENAQ 420
QY 421 QDKGIAIPIPIIDGESRVI IQVYDNGHQFQRPPTSPAPSPSPSVLERANDVLWLSITAAFY 480
DB 421 QDKGIAIPIPIIDGESRVI IQVYDNGHQFQRPPTSPAPSPSPSVLERANDVLWLSITAAFY 480
QY 481 PQSTYASSTGPPVVSNSVTLVNVATCAQAVASPLDWTVTYDGPPLSTIQGYSTFEVLP 540
DB 481 EQSTYASSTGPPVVSNSVTLVNVATCAQAVASPLDWTVTYDGPPLSTIQGYSTFEVLP 540
QY 541 LRGKLSFWEAGTICAGYPNYNTASDGLLVENAACTHEVAISTYITSLICAPVSIKAVAV 600
DB 541 LRGKLSFWEAGTICAGYPNYNTASDGLLVENAACTHEVAISTYITSLICAPVSIKAVAV 600
QY 601 LARHSALALGDLILYAPAKARTFDGPPPPPIIAGCAPSPSVAPAGPGRKMKVSKTPRI 660
DB 601 LARHSALALGDLILYAPAKARTFDGPPPPPIIAGCAPSPSVAPAGPGRKMKVSKTPRI 660

```


1 PRIOR APPLICATION DATA:
2 APPLICATION NUMBER: 08/542,644
3 FILING DATE: <unknown>
4 ATTORNEY/AGENT INFORMATION:
5 NAME: Fabian, Gary R.
6 REGISTRATION NUMBER: 33,875
7 REFERENCE/CKPT NUMBR: 4600-0294.30
8 TELECOMMUNICATION INFORMATION:
9 TELEPHONE: (415) 324-0880
10 TELEFAX: (415) 324-0960
11 INFORMATION FOR SEQ ID NO: 25:
12 SEQUENCE CHARACTERISTICS:
13 LENGTH: 540 amino acids
14 TYPE: amino acid
15 STRANDEDNESS: Hepatitis E virus (Burma strain)
16 TOPOLOGY: linear
17 MOLECULE TYPE: protein
18 HYPOTHETICAL: NO
19 ORIGINAL SOURCE:
20 SEQUENCE DESCRIPTION: SEQ ID NO: 25:
21 US-09-769-066-25

Query Match 81.8% Score 540; DB 10; Length 540;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 AVAPAHDTTPVPVIVDSRGAILLRQYNLSTPLTSSVATGNNLVLYAAPLPLPLQDGTN 171
Db 1 AVAPAHDTTPVPVIVDSRGAILLRQYNLSTPLTSSVATGNNLVLYAAPLPLPLQDGTN 60
QY 172 THIMATEASNYAQRVARATIRYRPLVNAVGGYATISFWPQTTTTPSVIMNSITSD 231
Db 61 THIMATEASNYAQRVARATIRYRPLVNAVGGYATISFWPQTTTTPSVIMNSITSD 120
QY 232 VRILVQPGIASLVIVSEKLVHYNQWRSVETSGVAEEATSGLVMLC IHGSLVNSYNT 291
Db 121 VRILVQPGIASLVIVSEKLVHYNQWRSVETSGVAEEATSGLVMLC IHGSLVNSYNT 180
QY 292 PYTCALGALIDFALIFEPNNITPVSYSSTAPRPLPGAGCJAFLLITTAIKPMK 351
Db 181 PYTCALGALIDFALIFEPNNITPVSYSSTAPRPLPGAGCJAFLLITTAIKPMK 240
QY 352 DLYPTSTNGVCFICGCIATLFLNLADTLIGLPTLISAGQGLPSPPVVSANGPTVK 411
Db 241 DLYPTSTNGVCFICGCIATLFLNLADTLIGLPTLISAGQGLPSPPVVSANGPTVK 360
QY 411 LYSVSNAGQGLPSPPVVSANGPTVKLPGAGCJAFLLITTAIKPMK 471
Db 361 LYSVSNAGQGLPSPPVVSANGPTVKLPGAGCJAFLLITTAIKPMK 320
QY 472 WLSLTAAPYQSTYSGSTFVVSVSQVTVVNVATCAQAVRSFWTVVTLDPFNSTIQQ 531
Db 321 WLSLTAAPYQSTYSGSTFVVSVSQVTVVNVATCAQAVRSFWTVVTLDPFNSTIQQ 420
QY 532 YKSTFVWLEPCVSPWIACTKACQVYVYNTASQQLIVENACHPVAISTYTSIGAC 591
Db 421 YKSTFVWLEPCVSPWIACTKACQVYVYNTASQQLIVENACHPVAISTYTSIGAC 480
QY 592 PVSISAVAVLAPHSALALELIDLYAPARATFQDQCNEDRDLGCGCAFQSTVAELKLE 651
Db 481 PVSISAVAVLAPHSALALELIDLYAPARATFQDQCNEDRDLGCGCAFQSTVAELKLE 640

RESULT 5
US-09-769-066-25
Sequence 27 Application us-09-769-066
Patent Ref. US20020307360A1
GENERAL INFORMATION:
Applicant: Fuest, Thomas P
McAtee, C. Patrick
Yarbough, Patrice O.
Zhang, Yifan

1 TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
2 NUMBER OF SEQUENCES: 41
3 CORRESPONDENCE ADDRESS:
4 ADDRESSEE: Dehlinger & Associates
5 STREET: 350 Cambridge Ave., Suite 250
6 CITY: Palo Alto
7 STATE: CA
8 COUNTRY: USA
9 ZIP: 94306
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE: floppy disk
12 COMPUTER: IBM PC compatible
13 OPERATING SYSTEM: PC-BUS/MS-DOS
14 SOFTWARE: Patent In Release #1.0, Version #1.25
15 CURRENT APPLICATION DATA: 08/542,644
16 FILING DATE: <unknown>
17 ATTORNEY/AGENT INFORMATION:
18 NAME: Fabian, Gary R.
19 REGISTRATION NUMBER: 33,875
20 REFERENCE/CKPT NUMBR: 4600-0294.40
21 TELECOMMUNICATION INFORMATION:
22 TELEPHONE: (415) 324-0880
23 TELEFAX: (415) 324-0960
24 INFORMATION FOR SEQ ID NO: 27:
25 SEQUENCE CHARACTERISTICS:
26 LENGTH: 525 amino acids
27 TYPE: amino acid
28 STRANDEDNESS: Hepatitis E virus (Burma strain)
29 TOPOLOGY: linear
30 MOLECULE TYPE: protein
31 HYPOTHETICAL: NO
32 ORIGINAL SOURCE:
33 SEQUENCE DESCRIPTION: SEQ ID NO: 27:
34 US-09-769-066-27

Query Match 79.5% Score 525; DB 10; Length 525;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 AVAPAHDTTPVPVIVDSRGAILLRQYNLSTPLTSSVATGNNLVLYAAPLPLPLQDGTN 171
Db 1 AVAPAHDTTPVPVIVDSRGAILLRQYNLSTPLTSSVATGNNLVLYAAPLPLPLQDGTN 60
QY 172 THIMATEASNYAQRVARATIRYRPLVNAVGGYATISFWPQTTTTPSVIMNSITSD 231
Db 61 THIMATEASNYAQRVARATIRYRPLVNAVGGYATISFWPQTTTTPSVIMNSITSD 120
QY 232 VRILVQPGIASLVIVSEKLVHYNQWRSVETSGVAEEATSGLVMLC IHGSLVNSYNT 291
Db 121 VRILVQPGIASLVIVSEKLVHYNQWRSVETSGVAEEATSGLVMLC IHGSLVNSYNT 180
QY 292 PYTCALGALIDFALIFEPNNITPVSYSSTAPRPLPGAGCJAFLLITTAIKPMK 351
Db 181 PYTCALGALIDFALIFEPNNITPVSYSSTAPRPLPGAGCJAFLLITTAIKPMK 240
QY 352 DLYPTSTNGVCFICGCIATLFLNLADTLIGLPTLISAGQGLPSPPVVSANGPTVK 411
Db 241 DLYPTSTNGVCFICGCIATLFLNLADTLIGLPTLISAGQGLPSPPVVSANGPTVK 360
QY 411 LYSVSNAGQGLPSPPVVSANGPTVKLPGAGCJAFLLITTAIKPMK 471
Db 361 LYSVSNAGQGLPSPPVVSANGPTVKLPGAGCJAFLLITTAIKPMK 320
QY 472 WLSLTAAPYQSTYSGSTFVVSVSQVTVVNVATCAQAVRSFWTVVTLDPFNSTIQQ 531
Db 321 WLSLTAAPYQSTYSGSTFVVSVSQVTVVNVATCAQAVRSFWTVVTLDPFNSTIQQ 420

QY 224 MNSITSTVDRIIVQVGIASELVIPSERIHYRNOGMRSVETSGVAEEATSLVLMCLHGS 283
 113 MNSITSTVDRIIVQVGIASELVIPSERIHYRNOGMRSVETSGVAEEATSLVLMCLHGS 172

RESULT 8

US-09-769-066-26
 : Sequence 26, Application US/09769065
 : Patent No. US20020107360A1
 : GENERAL INFORMATION:
 : APPLICANT: Fuerst, Thomas R.
 : McAlee, C. Patrick
 : Yarbough, Patrice O.
 : Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: 450 Cambridge Ave., Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09-769-066
 FILING DATE: 24-Jan-2001
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/542,634
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Fabian, Gary R.
 REGISTRATION NUMBER: 33,875
 REFERENCE/EXETER NUMBER: 4600-0293.30
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 26:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 540 amino acids
 TYPE: amino acid
 STRANDEDNESS: Hepatitis E virus (Mexico strain)
 TOPOLOGY: linear

MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 SEQUENCE DESCRIPTION: SEQ ID NO: 26:
 US-09-769-066-26

Query Match 18.2% Score 120; DB 10; Length 540;
 Best Local Similarity 100.0%; Pred. No. 2.2e-101;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 LPIQDQTNTHIMATASNYAQPVAATIPYPPVPAVAGGYAISFWPQTITPTPSVD 223
 113 LPIQDQTNTHIMATASNYAQPVAATIPYPPVPAVAGGYAISFWPQTITPTPSVD 112

RESULT 9

US-09-769-066-16
 : Sequence 16, Application US/09769066
 : Patent No. US20020107360A1
 : GENERAL INFORMATION:
 : APPLICANT: Fuerst, Thomas R.
 : McAlee, C. Patrick
 : Yarbough, Patrice O.
 : Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: 450 Cambridge Ave., Suite 250
 CITY: Palo Alto

APPLICANT: Fuerst, Thomas R.
 McAlee, C. Patrick
 Yarbough, Patrice O.
 Zhang, Yifan
 TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: 450 Cambridge Ave., Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09-769-066
 FILING DATE: 24-Jan-2001
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/542,634
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Fabian, Gary R.
 REGISTRATION NUMBER: 33,875
 REFERENCE/EXETER NUMBER: 4600-0293.30
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 549 amino acids
 TYPE: amino acid
 STRANDEDNESS: Hepatitis E virus (Mexico strain)
 TOPOLOGY: linear

MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 SEQUENCE DESCRIPTION: SEQ ID NO: 16:
 US-09-769-066-16

Query Match 18.2% Score 120; DB 10; Length 549;
 Best Local Similarity 100.0%; Pred. No. 2.2e-101;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 LPIQDQTNTHIMATASNYAQPVAATIPYPPVPAVAGGYAISFWPQTITPTPSVD 223
 113 LPIQDQTNTHIMATASNYAQPVAATIPYPPVPAVAGGYAISFWPQTITPTPSVD 112

QY 224 MNSITSTVDRIIVQVGIASELVIPSERIHYRNOGMRSVETSGVAEEATSLVLMCLHGS 283
 113 MNSITSTVDRIIVQVGIASELVIPSERIHYRNOGMRSVETSGVAEEATSLVLMCLHGS 172

RESULT 10

US-09-769-066-14
 : Sequence 14, Application US/09769066
 : Patent No. US20020107360A1
 : GENERAL INFORMATION:
 : APPLICANT: Fuerst, Thomas R.
 : McAlee, C. Patrick
 : Yarbough, Patrice O.
 : Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: 450 Cambridge Ave., Suite 250
 CITY: Palo Alto

```

1  STATE: CA
2  COUNTRY: USA
3  ZIP: 94306
4  COMPUTER RELEASABLE FORM:
5  MEDIUM TYPE: Floppy disk
6  COMPUTER: IBM PC compatible
7  OPERATING SYSTEM: PC DOS/MS-DOS
8  SOFTWARE: Patent In Release #1.0, Version #1.25
9  CURRENT APPLICATION DATA:
10 APPLICATION NUMBER: US2007669,066
11 FILING DATE: 24-Jan-2001
12 CLASSIFICATION: -BULK9900-
13
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: 08/542,634
16 FILING DATE: UNKNOWN
17 ATTORNEY/AGENT INFORMATION:
18 NAME: Patton, Gary R.
19 PESTIFICATION NUMBER: 33,075
20 REFERENCE/CHECKET NUMBER: 4609,0294,30
21 TELECOMMUNICATION INFORMATION:
22 TELEPHONE: (415) 424,0860
23 TELEFAX: (415) 424,0960
24 INFORMATION FOR SEQ ID NO: 14:
25     SEQUENCE CHARACTERISTICS:
26         LENGTH: 660 amino acids
27         TYPE: amino acid
28     SIMILARITY: Hepatitis E Virus (Mexico Strain)
29     topology: open-2
30
31     topology: linear
32
33     Molecule Type: protein
34     HYDROPHILIC: NO
35     ORIGINAL SOURCE:
36     SEQUENCE DESCRIPTION: SEQ ID NO: 14:
37     309 360 066,14

```

[illegible]

```

1  RELEASE 11
2  US 09 408 117 92
3  Sequence 92, Application US/0946147A
4  Publication No. US20090049601A1
5  GENERAL INFORMATION:
6  APPLICANT: Abbott Laboratories
7  APPLICANT: Scherander, George G.
8  APPLICANT: Erker, James L.
9  APPLICANT: Desai, Surresh M.
10 APPLICANT: Dawson, George J.
11 APPLICANT: Moshabian, L. K.
12 TITLE OF INVENTION: METHODS AND COMPOSITIONS
13 TITLE OF INVENTION: HEPATITIS VIRUS
14 FILE REFERENCE: 6,232 US, P1
15 CURRENT APPLICATION NUMBER: US/09/468,147A
16 CURRENT FILING DATE: 1999-12-21
17 EARLIER APPLICATION NUMBER: US 06/176,141
18 EARLIER FILING DATE: 1998-10-15
19 EARLIER APPLICATION NUMBER: US 60/061,199
20 EARLIER FILING DATE: 1997-10-15
21 NUMBER OF SEQ ID NOS: 258
22 SOFTWARE: FASTSEQ for Windows Version 4.0
23 SEQ ID NO 92
24 LENGTH: 600
25 TYPE: FPI

```

ORGANISM: Hepatitis E Virus	FEATURE:	DESCRIPTION:
OTHER INFORMAL FN: X-04	00_Kn0506	at pgs:1 to 6: 484
OTHER INFORMAL FN: X-04	00_Kn0506	at pgs:1 to 6: 542
OTHER INFORMAL FN: X-04	00_Kn0506	at pgs:1 to 6: 595

DS-007-4108 147-92

Query Match	11.4%	Search	11.4%	Length	100%
Best Local Similarity	100.0%	100.0%	100.0%	4	6
Matches	75	Consistent	100.0%	Matches	100

```

RESULT: 12
SDS 09-468-147-175.
Sequence 175, Application US/0946817A
Publication No. US/0946817A1
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Schlabach, George J.
APPLICANT: Erick, James C.
APPLICANT: Desai, Suresh M.
APPLICANT: Dawson, George J.
APPLICANT: Mushawar, L. K.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATING
FILE REFERENCE: 6,828,851
CURRENT APPLICATION NUMBER: US/0946817A
CURRENT FILING DATE: 1999-12-21
EARLIER APPLICATION NUMBER: US/09174,141
EARLIER FILING DATE: 1998-10-15
EARLIER APPLICATION NUMBER: US/09096,199
EARLIER FILING DATE: 1997-10-15
NUMBER OF SEQ IDS NOS: 258
SOFTWARE: FastSeq for Windows Version 8.0
SEQ ID NO 175
LENGTH: 427
TYPE: PRI
ORGANISM: Homo sapiens E VIRUS
FEATURE:
OTHER INFORMATION: US 1 SEQ
OTHER INFORMATION: Xaa Back seq of cholest
OTHER INFORMATION: Xaa Back seq of cholest
OTHER INFORMATION: Xaa Back seq of cholest
SDS 09-468-147-175.

```

[illegible]

RESULT 13
US: 09-468 147 176
Sequence 176; Application ID: 090468147A
Publication No. US2008004961A1
GENERAL INFORMATION
APPLICANT: Abbott Laboratories
APPLICANT: Schindler, Theodore

: APPLICANT: Erker, James C.
 : APPLICANT: Desai, Suresh M.
 : APPLICANT: Dawson, George J.
 : APPLICANT: Mushahwar, I. K.
 : TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
 : TITLE OF INVENTION: HEPATITIS E VIRUS
 : FILE REFERENCE: 6232.US.P1
 : CURRENT APPLICATION NUMBER: US/09/468,147A
 : CURRENT FILING DATE: 1999-12-21
 : EARLIER APPLICATION NUMBER: US 09/173,141
 : EARLIER FILING DATE: 1998-10-15
 : EARLIER APPLICATION NUMBER: US 60/061,199
 : EARLIER FILING DATE: 1997-10-15
 : NUMBER OF SEQ ID NOS: 258
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 176
 : LENGTH: 427
 : TYPE: PRT
 : ORGANISM: Hepatitis E Virus
 : FEATURE:
 : OTHER INFORMATION: US-2 SC3
 : OTHER INFORMATION: Xaa - unknown or other at position 114
 : US-09-468-147-176

Query Match: 9.2%; Score 61; DB 11; Length 338;
 Best Local Similarity 100.0%; Pred. No. 1.8e-47;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 365 GCGTATLTNLAHTLGGHPTLTLSSAGGQFYSPVVSANPTVKIYTSVFNAAQQQFG 424

DB 32 GCGTATLTNLAHTLGGHPTLTLSSAGGQFYSPVVSANPTVKIYTSVFNAAQQQFG 91

QY 425 1 425

DB 92 1 92

RESULT 14

US-09-468-147-199

: Sequence 199, Application US/09/468,147A

: Publication No. US20030049601A1

: GENERAL INFORMATION:

: APPLICANT: Abbott Laboratories

: APPLICANT: Schlauder, George G.

: APPLICANT: Erker, James C.

: APPLICANT: Desai, Suresh M.

: APPLICANT: Dawson, George J.

: APPLICANT: Mushahwar, I. K.

: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING

: TITLE OF INVENTION: HEPATITIS E VIRUS

: FILE REFERENCE: 6232.US.P1

: CURRENT APPLICATION NUMBER: US/09/468,147A

: CURRENT FILING DATE: 1999-12-21

: EARLIER APPLICATION NUMBER: US 09/173,141

: EARLIER FILING DATE: 1998-10-15

: EARLIER APPLICATION NUMBER: US 60/061,199

: EARLIER FILING DATE: 1997-10-15

: NUMBER OF SEQ ID NOS: 258

: SOFTWARE: FastSeq for Windows Version 3.0

: SEQ ID NO 199

: LENGTH: 438

: TYPE: PRT

: ORGANISM: Hepatitis E Virus

: FEATURE:

: OTHER INFORMATION: cksorf2m-2.pep

US-09-468-147-199

Query Match: 9.2%; Score 61; DB 11; Length 338;

Best Local Similarity 100.0%; Pred. No. 1.8e-47;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 365 GCGTATLTNLAHTLGGHPTLTLSSAGGQFYSPVVSANPTVKIYTSVFNAAQQQFG 424

DB 32 GCGTATLTNLAHTLGGHPTLTLSSAGGQFYSPVVSANPTVKIYTSVFNAAQQQFG 91

DB 35 GCGTATLTNLAHTLGGHPTLTLSSAGGQFYSPVVSANPTVKIYTSVFNAAQQQFG 94

QY 425 1 425

DB 95 1 95

RESULT 15

US-09-468-147-200

: Sequence 200, Application US/09/468,147A

: Publication No. US20030049601A1

: GENERAL INFORMATION:

: APPLICANT: Abbott Laboratories

: APPLICANT: Schlauder, George G.

: APPLICANT: Erker, James C.

: APPLICANT: Desai, Suresh M.

: APPLICANT: Dawson, George J.

: APPLICANT: Mushahwar, I. K.

: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING

: TITLE OF INVENTION: HEPATITIS E VIRUS

: FILE REFERENCE: 6232.US.P1

: CURRENT APPLICATION NUMBER: US/09/468,147A

: CURRENT FILING DATE: 1999-12-21

: EARLIER APPLICATION NUMBER: US 09/173,141

: EARLIER FILING DATE: 1998-10-15

: EARLIER APPLICATION NUMBER: US 60/061,199

: EARLIER FILING DATE: 1997-10-15

: NUMBER OF SEQ ID NOS: 258

: SOFTWARE: FastSeq for Windows Version 3.0

: SEQ ID NO 200

: LENGTH: 438

: TYPE: PRT

: ORGANISM: Hepatitis E Virus

: FEATURE:

: OTHER INFORMATION: Plot2.2.6.pep

US-09-468-147-200

Query Match

Best Local Similarity 100.0%; Pred. No. 1.8e-47;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 365 GCGTATLTNLAHTLGGHPTLTLSSAGGQFYSPVVSANPTVKIYTSVFNAAQQQFG 424

DB 35 GCGTATLTNLAHTLGGHPTLTLSSAGGQFYSPVVSANPTVKIYTSVFNAAQQQFG 94

QY 425 1 425

DB 95 1 95

Search completed: August 5, 2003, 09:45:09

Job time: 28 secs

QY 421 QOKGIAIPHDIDGESVWVIGYDNGHQEDRPTSPAPSPFVLRANDVWLILTAABY 480
 DB 421 QOKGIAIPHDIDGESVWVIGYDNGHQEDRPTSPAPSPFVLRANDVWLILTAABY 480
 QY 481 DQSTVGSSTGVVVSNSVITVNVATGAQAVAPSIQWTKVTLGGPPLSTIQQYSPFFVLP 540
 DB 481 DQSTVGSSTGVVVSNSVITVNVATGAQAVAPSIQWTKVTLGGPPLSTIQQYSPFFVLP 540
 QY 541 LRGLSLWEAGTAKGYPYNYNTASDLILVRNAAGHRVAISTYTTTSIGAGVPSISAVAV 600
 DB 541 LRGLSLWEAGTAKGYPYNYNTASDLILVRNAAGHRVAISTYTTTSIGAGVPSISAVAV 600
 QY 601 LAPHSAIAGLFDLIDYAPAHITFDGRCPECPHPLGLGSAFQSIIVAEILQPLKMKVCKLDEL 660
 DB 601 LAPHSAIAGLFDLIDYAPAHITFDGRCPECPHPLGLGSAFQSIIVAEILQPLKMKVCKLDEL 660

RESULT 2

B44212

structural protein 2 precursor - hepatitis E virus (strain Mexico)

C:Species: hepatitis E virus

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999

C:Accession: B44212

R:Buang, C.C.; Nguyen, D.; Fernandez, J.; Yan, K.Y.; Fry, K.E.; Bradley, D.W.; Tam, A.W.

Virol. 191, 550-558, 1992

A:Title: Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HE

E-23-62) from a structural protein 2 status predicted

A:Reference number: A44212; MUID:93070867; PMID:1448913

A:Accession: B44212

A:Molecule type: genomic RNA

A:Residues: 1-659 <HDA>

A:Cross-references: GR-M74506; NID:q330017; PIRN:AAA45732.1; PID:q330020

C:Superfamily: hepatitis E virus structural protein 2

C:Keywords: structural protein

E-1-22-De-mex: Structural Sequences Status Predicted

E-23-62-Pr-Mex: Structural Protein 2 Status Predicted

Query Match 12 29; Score 8; DB 2; Length 1091

Best Local Similarity 100.0%; Pred. No. 8;

Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 LPLQGNTHIMAFASNAQYQVAVATIPPLVPRVAGGVALSIGWQTITPTISVD 223
 DB 154 LPLQGNTHIMAFASNAQYQVAVATIPPLVPRVAGGVALSIGWQTITPTISVD 223
 QY 224 MNSITSTDPVILVQGLIASPLVTPSPFLHYRNGWPSVETSGVAFPAISLVMGLHGS 283
 DB 224 MNSITSTDPVILVQGLIASPLVTPSPFLHYRNGWPSVETSGVAFPAISLVMGLHGS 283

RESULT 3

F83928

hypothetical protein B44212 [imported] Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: F83928

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512592; PMID:11058132

A:Accession: F83928

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 11091 <STO>

A:Cross-references: GR-AP001534; GR-BA000004; NID:g10174613; PIRN:PAK05949.1; GSPIR-GN00

C:Experimental source: strain C-125

C:Genetics:

A:Gene: B42330

Query Match 1.4%; Score 9; DB 2; Length 1091;
 Best Local Similarity 100.0%; Pred. No. 8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LLLLLLMFLP 15

DB 12 LLLLLLMFLP 20
 RESULT 4
 B58208
 F:Ramirez, J. Black fly snake (Incubated)
 C:Species: Elaphe obsoleta quadrivittata
 C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 07-May-1999
 C:Accession: B58208
 R:Rand, J.C.; Kasinaky, H.P.; Elsey, R.M.; Wright, C.L.; Rice, F.; Rolly, J.F.; Sharif,
 J. Biol. Chem. 271, 23547-23557, 1996
 A:Title: Protamines of reptiles.
 A:Reference number: A58208; MUID:96494458; PMID:8798564
 A:Accession: B58208
 A>Status: Preliminary
 A:Molecule type: protein
 A:Residues: 141 <REN>
 C:Superfamily: sperm histone

Query Match 1.2%; Score 8; DB 2; Length 41;
 Best Local Similarity 100.0%; Pred. No. 4.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 LREERERK 43

DB 13 SCPEPEEP 20

RESULT 5

F87649

Exho/Tail family protein [imported] Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 29-Apr-2001 #sequence_revision 29-Apr-2001 #text_change 20-Apr-2001

C:Accession: F87649

R:Norman, W.C.; Feidblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.P.; Heidelberg,
 K.; Jack, M.T.; Fields, E.T.; Laskin, E.S.; Parkin, A.S.; Glick, M.L.; Holt, C.H.; Po

n, J.; Prmelova, M.; White, O.; Salzberg, S.L.; Shapiro, J.; Venter, J.C.; Fraser, C

Proc. Natl. Acad. Sci. U.S.A. 98, 4140-4143, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87649; MUID:21172699; PMID:11250647

A:Accession: F87649

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 172 <GPO>

A:Cross-references: GR-AP005673; NID:41342416; PIRN:AAK2194.1; GSPIR-GN00148

C:Genetics:

A:Gene: GS342

Query Match

1.2%; Score 8; DB 2; Length 172;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 QSRGGRSS 44

DB 16 QSRGGRSS 24

RESULT 6

S05509

Phosphatase 11 cytochrome b161-1 p16 protein 4 precursor - Chlamydomonas reinhardtii

A:Reference number: S05509

C:Species: Chlamydomonas reinhardtii

C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Sep-1997

C:Accession: S05509

R:Mayfield, S.P.; Schirmer-Rahire, M.; Frank, G.; Zuber, H.; Rechbach, J.L.

Plant Mol. Biol. 12, 683-693, 1989

A:Title: Analysis of the genes of the cbe1 and cbe3 proteins of the photosystem II co

A:Reference number: S05509

A:Accession: S05509

A:Molecule type: DNA

A:Residues: 139 <NAY>

A:Cross-references: EMBL:X13822; NID:419154; PIR:q418164

Cited from:

A:Gene: pds4
 C:Keywords: chloroplast; membrane protein; photosynthesis; photosystem II
 C:Species: Arabidopsis thaliana
 C:Date: 04-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 21-Jul-2000
 C:Accession: S72224

Query Match 1.2% Score 8; DB 2; Length 199;
 Best local similarity 100.0%; Pred. No. 17;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 93 RFAVASRP 100
 ID 11111111
 ID 11 RFAVASRP 17

RESULT 7

S72224
 C:Keywords: chloroplast; membrane protein; precursor
 C:Species: Arabidopsis thaliana
 C:Date: 04-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 21-Jul-2000
 C:Accession: S72224

Query Match 1.2% Score 8; DB 2; Length 199;
 Best local similarity 100.0%; Pred. No. 17;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 93 RFAVASRP 100
 ID 11111111
 ID 11 RFAVASRP 17

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 17;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 93 RFAVASRP 100
 ID 11111111
 ID 11 RFAVASRP 17

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 17;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 93 RFAVASRP 100
 ID 11111111
 ID 11 RFAVASRP 17

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 17;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 93 RFAVASRP 100
 ID 11111111
 ID 11 RFAVASRP 17

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 17;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 93 RFAVASRP 100
 ID 11111111
 ID 11 RFAVASRP 17

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 17;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 93 RFAVASRP 100
 ID 11111111
 ID 11 RFAVASRP 17

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 17;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 93 RFAVASRP 100
 ID 11111111
 ID 11 RFAVASRP 17

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 17;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 93 RFAVASRP 100
 ID 11111111
 ID 11 RFAVASRP 17

WIDEH

DLA protein - human herpesvirus 2 (strain Hb22)
 C:Species: Human herpesvirus 2
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: J01496

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 20;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 141 SPTSSVA 148
 ID 11111111
 ID 75 SPTSSVA 82

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 20;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 141 SPTSSVA 148
 ID 11111111
 ID 75 SPTSSVA 82

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 20;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 141 SPTSSVA 148
 ID 11111111
 ID 75 SPTSSVA 82

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 20;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 141 SPTSSVA 148
 ID 11111111
 ID 75 SPTSSVA 82

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 20;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 141 SPTSSVA 148
 ID 11111111
 ID 75 SPTSSVA 82

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 20;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 141 SPTSSVA 148
 ID 11111111
 ID 75 SPTSSVA 82

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 20;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 141 SPTSSVA 148
 ID 11111111
 ID 75 SPTSSVA 82

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 20;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 141 SPTSSVA 148
 ID 11111111
 ID 75 SPTSSVA 82

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 20;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 141 SPTSSVA 148
 ID 11111111
 ID 75 SPTSSVA 82

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 20;
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 141 SPTSSVA 148
 ID 11111111
 ID 75 SPTSSVA 82

A:Gene: rps2
 A:Genome: chloroplast
 A:Superfamily: Escherichia coli ribosomal protein S2
 C:Keywords: ribosomal protein, protein synthesis, ribosome

Query Match 1.2% Score 8; DB 2; Length 241;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 324 PTE1SSA 321
 |||||
 DB 231 PTE1SSA 238

RESULT 12
 145737
 Hypothetical protein F24M12.100 Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 04-Feb-2000 #sequence_revision: 64 Feb-2000 #tx_change: 64 Feb-2000
 C:Accession: T45737
 R:Vitale, P.; Liguori, R.; Flores, M.; Atgikou, A.; De Simone, V.; Mewes, H.W.; Lemcke, A. Submitted to the Protein Sequence Database, December 1999
 A:Reference number: T23012
 A:Accession: T45737
 A:Status: preliminary
 A:Molecule type: RNA
 A:Residues: 1-252 <VT>
 A:Cross-references: EMBL:AL132980
 A:Experimental source: Cultivar Columbia; BAC clone F24M12
 C:Genetics:
 A:Map position: 3
 A:Introns: 232/1
 A:Note: F24M12.100

Query Match 1.2% Score 8; DB 2; Length 252;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 RSGSGGG 40
 |||||
 DB 131 RSGSGGG 138

RESULT 13
 148649
 Hypothetical protein T15N1.150 Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse ear cress)
 C:Date: 20-Apr-2000 #sequence_revision: 70 Apr-2000 #tx_change: 20 Apr-2000
 C:Accession: T48649
 R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, J.; De
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: T24493
 A:Accession: T48649
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-273 <BEV>
 A:Cross-references: EMBL:AL163792
 A:Experimental source: Cultivar Columbia; BAC clone T15N1
 C:Genetics:
 A:Map position: 5
 A:Introns: 72/1; 118/2; 123/3; 219/3; 340/3
 A:Note: T15N1.150

Query Match 1.2% Score 8; DB 2; Length 273;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 ISPLTSSV 147
 |||||
 DB 48 ISPLTSSV 55

RESULT 14

A87184
 thiosulfate sulfurtransferase [imported] - Mycobacterium leprae
 C:Species: Mycobacterium leprae
 C:Date: 20 Apr 2001 #sequence_revision: 20 Apr 2001 #tx_change: 17 May-2002
 C:Accession: A87184
 R:Cole, S.V.; Chalmers, K.; Parkhill, J.; James, K.D.; Thomas, N.; Kierstead, P.; Kierstead, S.M.; Devlin, K.; Bullock, S.; Fellgett, L.; Fraser, A.; Harkin, N.; Harkin, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A:Author: Fellgett, S.; Cooper, K.; James, K.D.; Thomas, N.; Kierstead, P.; Kierstead, S.M.; Devlin, K.; Bullock, S.; Fellgett, L.; Fraser, A.; Harkin, N.; Harkin, M.A.; Rutherford, K.M.
 A:Title: Massive gene decay in the leprosy bacillus
 A:Reference number: A95099; PMID:1128742; EMBL:1124092
 A:Accession: A87184
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-277 <STP>
 A:Cross-references: DB:AL145082; EMBL:CA31153.1; GSPRIGNO:147
 C:Genetics:
 A:Gene: cysA3
 C:Superfamily: thiosulfate sulfurtransferase

Query Match 1.2% Score 8; DB 2; Length 277;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 521 LDKRPST 528
 |||||
 DB 118 LDKRPST 125

RESULT 15
 H70413
 L-alanine synthase related protein cysW - Apifex aceticus
 C:Species: Apifex aceticus
 C:Date: 09 May 1998 #sequence_revision: 08 May 1998 #tx_change: 17-Mar-2000
 C:Accession: H70413
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Jones, A.L.; Graham, D.E.; V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Apifex aceticus
 A:Reference number: A70409; PMID:9819666; PMID:9537420
 A:Accession: H70413
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-292 <APF>
 A:Cross-references: DB:AF000675; NID:42382863; FID:AA00494.1; FID:4282874; GB:AE00
 A:Experimental source: strain VF5
 C:Genetics:
 A:Gene: cysW
 C:Superfamily: cysW protein

Query Match 1.2% Score 8; DB 2; Length 292;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 VETSGVAE 268
 |||||
 DB 90 VETSGVAE 97

Search completed: August 5, 2003 00:24:10
 Job time: 24 secs


```

QY 181 NYAGYVARATATVETFLVNAVGGVAISLSHWPTTTTTTTSVEMNSITSTVEILLVSGGI 240
DB 181 NYAGYVARATATVETFLVNAVGGVAISLSHWPTTTTTTTSVEMNSITSTVEILLVSGGI 240
QY 241 ASELVTPSEPLHPNCGWPSVETSGVAPPEATSGVLMCLHGSLVNSVNTPTPTGALGL 300
DB 241 ASELVTPSEPLHPNCGWPSVETSGVAPPEATSGVLMCLHGSLVNSVNTPTPTGALGL 300
QY 301 DFALELEFNLTPGNTNTVPSYSSTARHPLRPGAGTAEIITTAATRFMKDLFTSTNG 360
DB 301 DFALELEFNLTPGNTNTVPSYSSTARHPLRPGAGTAEIITTAATRFMKDLFTSTNG 360
QY 361 VGEIGETALTPNADTLGLPTLGLISSANQGLFSPVVSANGPTVKLYTSVENAQ 420
DB 361 VGEIGETALTPNADTLGLPTLGLISSANQGLFSPVVSANGPTVKLYTSVENAQ 420
QY 421 QDRGIALPHDIDLGESRVVLIQYDNGHEQDRTPSPAPSRPFSVLKANDVLMWLSLTAAY 480
DB 421 QDRGIALPHDIDLGESRVVLIQYDNGHEQDRTPSPAPSRPFSVLKANDVLMWLSLTAAY 480
QY 481 DOSTYSSSGPVVVSQSVTLVNVATGACAVAPSLDWTKVTLGRLPLSTIQOYSKTFVLP 540
DB 481 DOSTYSSSGPVVVSQSVTLVNVATGACAVAPSLDWTKVTLGRLPLSTIQOYSKTFVLP 540
QY 541 LRKLSFWPAGITKAGYPYNYNTASDGLLVENAGCHKVALSYTISLUGAGPWSISAVAV 600
DB 541 LRKLSFWPAGITKAGYPYNYNTASDGLLVENAGCHKVALSYTISLUGAGPWSISAVAV 600
QY 601 LAPHSALALDRTLDYAPARHTDDPCPCPLSHGAGCAFGSTVAELQPLMKVKTRREL 660
DB 601 LAPHSALALDRTLDYAPARHTDDPCPCPLSHGAGCAFGSTVAELQPLMKVKTRREL 660

RESULT 2
VST2_HEVPA STANDARD: PRT: 660 AA
AC P3426;
DT 01-FEB-1994 (Rel. 28, Created)
DI 01-FEB-1994 (Rel. 28, Last sequence update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Pakistan) (HEV).
OC Viruses; ssRNA positive strand viruses, no DNA stage.
CC Hepatitis E-like viruses.
CX NCBI_TaxID=33774;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92115700; PubMed=1711327;
RA Tsarev S.A., Emerson S.G., Paves G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RA "Characterization of a prototype strain of hepatitis E virus."
PI Proc Natl Acad Sci U S A 89:559-563(1992)
CC -1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSULATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/industry
CC or send an email to license@sib.ch).
CC
DB EMBL: M80581; AAA45727.1;
DR InterPro: IPR004261; SP2.
DR Pfam: PF03014; SP2; 1.
KW Signal.
FT CHAIN 1 22 BY SIMILARITY.
FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 660 AA; 70550 MW; 8065B553CFB4ED3 C0664.

```

```

Query Match 61.8%; Score 408; DB 1; Length 660;
Best Local Similarity 99.9%; Ref. No. 0;
Matches 508; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 97 ASRRRTTAAAPLJAVAPAHDTTPVVDSSRGALLRQYNLSITPTSSVAIGINIVLY 154
DB 97 ASRRRTTAAAPLJAVAPAHDTTPVVDSSRGALLRQYNLSITPTSSVAIGINIVLY 154
QY 157 AAPLSPLLPLOGTNTHIMATEASNYAYRVARATATVETFLVNAVGGVAISLSHWPT 214
DB 157 AAPLSPLLPLOGTNTHIMATEASNYAYRVARATATVETFLVNAVGGVAISLSHWPT 214
QY 217 TPTTSVDMNSITSTVDRILLVQPGTASHLIVTTSERLHYRNOGWSVETSVAEEATSLV 276
DB 217 TPTTSVDMNSITSTVDRILLVQPGTASHLIVTTSERLHYRNOGWSVETSVAEEATSLV 276
QY 277 MLCIHGSVNSYNTTYTCALGLIDFALDITNLTGTNTTRVSRYSSTARIHLRPGAG 346
DB 277 MLCIHGSVNSYNTTYTCALGLIDFALDITNLTGTNTTRVSRYSSTARIHLRPGAG 346
QY 337 GTAEIITTAATRFMKDLFTSTNGVCEIGRGIALTLNLAUTLGLPTLGLISSANQGL 396
DB 337 GTAEIITTAATRFMKDLFTSTNGVCEIGRGIALTLNLAUTLGLPTLGLISSANQGL 396
QY 397 YFVFWVADPTTPVLYTVETVADQGRKLAIPHEIDTTPVVDSSRGALLRQYNLSITPT 454
DB 397 YFVFWVADPTTPVLYTVETVADQGRKLAIPHEIDTTPVVDSSRGALLRQYNLSITPT 454
QY 457 ATSRRTNVLKANDVLMWLSLTAAYRQYRYSSTDPVAVSVSVTLVNVATGACAVAPSL 514
DB 457 ATSRRTNVLKANDVLMWLSLTAAYRQYRYSSTDPVAVSVSVTLVNVATGACAVAPSL 514
QY 517 TPTTSVDMNSITSTVDRILLVQPGTASHLIVTTSERLHYRNOGWSVETSVAEEATSLV 576
DB 517 TPTTSVDMNSITSTVDRILLVQPGTASHLIVTTSERLHYRNOGWSVETSVAEEATSLV 576
QY 577 HEVATITVTISLUGAGPWSISAVAVLAPHS 605
DB 577 HEVATITVTISLUGAGPWSISAVAVLAPHS 605

RESULT 4
VST2_HEVPA STANDARD: PRT: 660 AA.
AC Q04611;
DT 01-OCT-1993 (Rel. 27, Created)
DI 01-OCT-1993 (Rel. 27, Last sequence update)
DE Structural protein 2 precursor (ORF2).
OC Hepatitis E virus (strata Myanmar) (HEV).
CC Viruses; ssRNA positive strand viruses, no DNA stage.
CX Ref.133. E-like viruses.
RN [1]_taxid=41769;
RP SEQUENCE FROM N.A.
RX MEDLINE=93227574; PubMed 8470471;
RA Aye I.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
RA Rikihisa T., Winn K.;
RA "Sequence and gene structure of the hepatitis E virus isolated from
RA Myanmar."
RL Virus genes 7:95-109(1993).
CC -1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSULATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/industry
CC or send an email to license@sib.ch).
CC

```



```

RI photosystem II complex of Chlamydomonas reinhardtii."
RL Plant Mol. Biol. 12:493-493(1994).
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
CC with the photosystem II complex.
CC -1- SIMILARITY: BELONGS TO THE PSBO FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch)
CC or send an email to license@sib-sib.ch.
CC
DR EMBL: X13832, CAA32051.1.
DR PIR: S05509; S05509.
KW Photosynthesis; Photosystem II; Chloroplast; Transit peptide;
KW Thylakoid; Membrane.
FT TRANSIT 1 51 CHLOROPLAST.
FT CHAIN 52 190 OXYGEN-EVOLVING ENHANCER PROTEIN 3.
SQ SEQUENCE 190 AA: 21824 MW: 30707.734 kDa; cpe64;
Query Match 1.28; Score 8; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 6;
Matches R; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 93 PPAVASPP 100
DB 10 RPAVASRR 17
|||||
RESULT 7
ID PSBO-VOLCA STANDARD; PRT; 202 AA.
AC Q41643;
DT 16-OCT-2001 (Rel. 40, Created)
DI 16-OCT-2001 (Rel. 40, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oxygen-evolving enhancer protein in 3, chloroplast precursor (cpe64)
GN PSBO.
OS Volvox carteri.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales.
OC Volvocaceae; Volvox.
OX NCBI_TaxID=3067;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=f. Naqariensis / HK10;
RX MEDLINE=95376499; PubMed=8781179;
RA Choi G., Przybylska M., Straus D.;
RT "Three abundant germ line-specific transcripts in Volvox carteri
RI encode photosynthetic proteins."
RL Curr. Genet. 40:347-355(1996)
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
CC with the photosystem II complex.
CC -1- SIMILARITY: BELONGS TO THE PSBO FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch)
CC or send an email to license@sib-sib.ch.
CC
DR EMBL: U22330; AAR40980.1.
DR PIR: S72324; S72324.
DR TIGRGRAMS: TIGR01409; TAT_signal_seq: 1.
KW Photosynthesis; Photosystem II; Chloroplast; Transit peptide;
KW Thylakoid; Membrane.
FT TRANSIT 1 54 CHLOROPLAST (POTENTIAL).
FT CHAIN 55 202 OXYGEN-EVOLVING ENHANCER PROTEIN 3.
SQ SEQUENCE 202 AA: 22290 MW: 31033.038 kDa; cpe64;

```

```

Query Match 1.28; Score 8; DB 1; Length 202;
Best Local Similarity 100.0%; Pred. No. 6;
Matches R; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 93 RPAVASRR 100
DB 10 RPAVASRR 17
|||||
RESULT 8
ID OLI3-HSV2H STANDARD; PRT; 241 AA.
AC P28279;
DT 01-DEC-1992 (Rel. 24, Created)
DI 01-DEC-1992 (Rel. 24, Last sequence update)
DI 01-DEC-1992 (Rel. 24, Last annotation update)
DE Protein UL3.
GN UL3.
OS Herpes simplex virus (Type 2 / strain HG52).
OC Viruses; dsDNA viruses; dsDNA state; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10415;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=92114459; PubMed=1622067;
RA McGeech D.J., Cunningham C., McIntyre G., Dolan A.;
RT "Comparative sequence analysis of the long repeat regions and
RT adjoining parts of the long unique regions in the genomes of herpes
RT simplex viruses types 1 and 2."
RL J. Gen. Virol. 72:3057-3075(1991).
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL3,
CC HSV-2 UL3, EBV 1.60, AND VZV 58.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch)
CC or send an email to license@sib-sib.ch.
CC
DR EMBL: J01476; XAA 1266.1.
DR PIR: J01496; WMBHRK.
DR InterPro: IPR005045; Herpes_UL3.
DR Pfam: PF03369; Herpes_UL3; 1.
SQ SEQUENCE 243 AA: 25649 MW: 84961.908 kDa; cpe64;
Query Match 1.28; Score 8; DB 1; Length 243;
Best Local Similarity 100.0%; Pred. No. 7;
Matches R; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 141 SPLTSSVA 148
DB 75 SPLTSSVA 82
|||||
RESULT 9
ID P21496 STANDARD; PRT; 241 AA.
AC P51349;
DT 01-OCT-1996 (Rel. 34, Created)
DI 01-OCT-1996 (Rel. 34, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 60S ribosomal protein S2.
GN RPS2.
OS Porphyra purpurea.
OC Chloroplast.
OC Eukaryota; Eukaryota; Bandiophyta; Bandiophyceae; Porphyra.
OX NCBI_TaxID=27877;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport.
RA Keith M.E., Munholland J.;

```

RA "Complete nucleotide sequence of the porphyra purpurea chloroplast
 RA genome".
 RA Plant Mol. Biol. Rep. 14:343-345(1995).
 RA 1. SIMILARITY: RELINGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)
 CC -----
 RA EMBL: 048404; AAC08135.1; .
 DR PIR: S74179; S74179.
 DR IANAP: MF 00241; .
 DR InterPro: IPR001865; Ribosomal_S2.
 DR Pfam: PF00318; Ribosomal_S2.
 DR PRINTS: PR00395; RIBOSOMAL_S2.
 DR PROSITE: PS00945; RIBOSOMAL_S2.1; 1.
 DR PROSITE: PS00945; RIBOSOMAL_S2.2; 1.
 DR PROSITE: PS00945; RIBOSOMAL_S2.2; 1.
 KW Ribosomal protein; Chloroplast.
 SQ SEQUENCE 241 AA; 27198 MW; FA17DA96C3E7F0866 CRC64.

Query Match 1.28; Score 8; DB 1; Length 241;
 Best Local Similarity 100.00; Prod. No. 7.7;
 Matches 8; Conservative 0; Mismatches 0; Gaps 0.
 QY 004 PTELISSA 91
 DI 1111111
 DI 241 PTELISSA 248

 RES011 10
 DEFO APATH STANDARD; PRI: 274 AA.
 AC QY022; QY114;
 DI 28 FEB 2003 (rel. 41; Last sequence update)
 DI 28 FEB 2003 (rel. 41; Last annotation update)
 DE Peptide deformylase, chloroplast precursor (E035.1.80) (PDF)
 DE (Polypeptide deformylase).
 GN PDFB_BA15614660 [P15ML150].
 OS Arabidopsis thaliana (Mouse-ear cress).
 EC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 EC Spermatophyta; Magnoliopsida; Eudicotyledons; Core eudicots; Rosidae;
 EC Geraniales; Brassicales; Brassicaceae; Arabidopsida.
 OX NCBT TaxID 4702;
 RN 111
 RP SEQUENCE FROM N.A.
 RX MEDLINE 20514156; PubMed 11066042;
 RA Gattione C., Sorro A., Piore M., Boisson B., Meimel T.;
 RA "Identification of eukaryotic peptide deformylases reveals
 RA universality of N-terminal protein processing mechanisms".
 RL EMBL J. 19:5416-5429(2000).
 RN 121
 RF SEQUENCE FROM N.A.
 RC S1PAIN cv. Columbia;
 RX MEDLINE 21016721; PubMed 11140714;
 RA Tabata S., Kuroki T., Nakamura Y., Kato T., Asanizu E.,
 RA Miyajima H., Sasamoto S., Kikuta T., Hosonchi T., Kawashima K.,
 RA Kohari M., Matsuno M., Matsuo A., Muraki A., Nakayama S.,
 RA Kasezaki N., Narita K., Okazaki S., Shirai S., Takouchi T., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 RA Inoue E., Spillner L., Guo J., O'Shaughnessy A., Preston J.,
 RA Rothermund K., Murray J., Johnson D., Kohlring T., Nelson J.,
 RA Storking T., Popin K., Speith J., Sekhon M., Armstrong W., Dante M.,
 RA Jones H., Edwards J., Fryman J., Hadenstein B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozerovskiy P., Riley A., Strömatt G.,
 RA Warner M., Watson C., Wollam A., Yeakum M., Bell M., Dedhia N.,

RA Parnell L., Shah R., Redington M., Bonn S., Lee A., Baker J.,
 RA Kirschner E., Pelt E., Kint J., Baker A., Miller B., Maria M.,
 RA Mattenissen K., McComb W.P., Wells F.K., Murphy G., Rothermund K.,
 RA Vothkari R., Wambert R., Wambert R., Wambert R., Wambert R.,
 RA Enliam K., D., D., D., D., D., D., D., D., D., D., D., D., D.,
 RA Langham S.A., McNeillach P., Boddin J., Glynnepier W., Zimmermann W.,
 RA Ransperger U., Wodter H., Balke K., Wodter H., Wodter H.,
 RA van Schueren M., Pelt E., Kint J., Baker A., Miller B., Maria M.,
 RA Wollam M., Wollam M., Wollam M., Wollam M., Wollam M.,
 RA Feldpausch M., Lamberth S., Villard R., Gledits J., Ardiles W.,
 RA Wollam M., Wollam M., Wollam M., Wollam M., Wollam M.,
 RA Scheller C., Ziegler P., Wollam M., Wollam M., Wollam M.,
 RA "Sequence and analysis of chromosome 6 of the plant Arabidopsis
 RA thaliana".
 RA Nature 406: 825-826(2000).
 CC 1. POSITION: Resides the amino group from the N-terminal Met of
 CC newly synthesized proteins (by similarity).
 CC 1. CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H2O -> formate +
 CC methionyl peptide.
 CC 1. COFACTOR: Binds 1 iron(II) ion (by similarity).
 CC 1. SUBCELLULAR LOCATION: chloroplast (potential).
 CC 1. SIMILARITY: BELONGS TO THE PEPTIDYL DEFORMYLASE FAMILY.
 CC -----
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)
 CC -----
 RA EMBL: AF269165; AAC33980.1;
 DR EMBL: AL163792; CAB67633.1;
 DR HSP: P27251; 1DEF.
 DR InterPro: IPR000181; Pop_deformylase.
 DR Pfam: PF01327; Pop_deformylase.
 DR PRINTS: PR01576; POPDEFORMYLASE.
 DR Problem: P0004844; Pop_deformylase; 1.
 DR TIGRFAIMS: TIGR00079; pop_deformyl; 1.
 KW Protein biosynthesis; Hydrolase; Iron; chloroplast; Transit peptide.
 FT TRANSIT 1 2 274 CHLOROPLAST (PEPTIDYL).
 FT CHAIN 2 274 PEPTIDE DEFORMYLASE.
 FT METAL 171 171 IRON (BY SIMILARITY).
 FT METAL 213 213 IRON (BY SIMILARITY).
 FT ACT_SITE 214 214 IRON (BY SIMILARITY).
 FT METAL 217 217 IRON (BY SIMILARITY).
 FT COMBINED 295 295 S 6 (IN PDF_2).
 SQ SEQUENCE 273 AA; 30741 MW; 106A00E0A0000 (P004);

Query Match 1.28; Score 8; DB 1; Length 273;
 Best Local Similarity 100.00; Prod. No. 6.6;
 Matches 8; Conservative 0; Mismatches 0; Gaps 0.
 QY 140 TSPLISSW 147
 DI 1111111
 DI 48 TSPLISSW 55

 RES011 11
 DEFO MYCLE STANDARD; PRI: 277 AA.
 AC Q50046;
 DI 15 JUL 1998 (rel. 40; Created)
 DI 15 JUL 1998 (rel. 40; Last sequence update)
 DI 28 FEB 2003 (rel. 41; Last annotation update)
 DE Putative thiosulfate sulfotransferase (P0239.1.1) (Phodanese-like
 DE protein).
 GN CYSA OR CYSA3 OR M2198.
 OS Mycobacterium leprae.
 EC Bacterota; Acidobacteriota; Acidobacteriales; Actinomycetales;
 EC Corynebacteriales; Mycobacteriaceae; Mycobacterium.
 OX NCBT TaxID 1769;
 RN 111

```

RP SEQUENCE FROM N.A.
RA Smith D.R., Robinson K.;
RI Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RC STRAIN:TN;
RX MEDLINE=21126732; PubMed=11234002;
RA Cole S.T., Elguier K., Parkhill J., James K.D., Thomson N.P.,
RA Wheeler P.R., Honore N., Gattner T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor K.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hurrest T., Jarvis K., Jarvis K., Mearns J., Mearns S.,
RA Murphy J., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Butler S., Seeger K., Simon S., Simmonds M., Skellern J., Squares F.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Harrell B.G.;
RI "Massive gene decay in the leprosy bacillus.";
RL Nature 406:1007-1011(2003)
CC -!- FUNCTION: MAY BE A SULFOTRANSFERASE INVOLVED IN THE FORMATION OF
CC THIOSULFATE
CC -!- CATALYTIC ACTIVITY: Thiosulfate + cyanide = sulfite + thiocyanate.
CC -!- SIMILARITY: Contains 2 rhodanese domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC
CC EMBL: U15182; AAA62982.1;
DR EMBL: AL583924; CAC31153.1;
DR PIR: A87184; A87184.
DR Leprosoma; ML2198;
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR001307; Rhodanese
DR Pfam: PF00581; Rhodanese; 2.
DR SMART: SM00450; PHOD; 2.
DR PROSITE: PS00683; PHODANES_2; 1.
DR PROSITE: PS02066; PHODANES_3; 2.
KW Transferrase; Complete proteome; Repeat.
FT DOMAIN 18 125 RHODANES 1.
FT DOMAIN 154 274 RHODANES 2.
FT ACT_SITE 233 233 BY SIMILARITY.
SU SEQUENCE 277 AA; 31094 MW; 527E05C9FF93969D CPC64;

Query Match 1.28; Score 8; DB 1; Length 277;
Best Local Similarity 100.0%; Pred No R 7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 521 IIVPPLST 528
DB 118 LUCRPLST 125
IIIIIIII

RESULT 12
YNP1_CAEEL
ID YNP1_CAEEL STANDARD; PRT; 412 AA.
AC P34554;
DT 01-FEB-1994 (Rel. 28, Created)
DI 30-MAY-2000 (Rel. 39, Last sequence update)
DI 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein T05G5.1 in chromosome III precursor
GN T05G5.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chordata; Chordata; Phlebotomidae; Phlebotomidae;
OC Rhabditidae; Pelodidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN: Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;

```

```

KA Wilson R., Ainsworth R., Anderson K., Baynes C., Berks M.,
KA Benfield J., Burton J., Connell M., Copsey L., Cooper J., Coulson A.,
KA Craxton M., Dear S., De Z., Durbin P., Pavolli A., Fraser A.,
KA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jior M.,
KA Johnston L., Jones M., Kershaw J., Kirsten J., Laisner N.,
KA Lathille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
KA Parsons J., Pary G., Pickett L., Pickett A., Sandberg R., Shownkeen R.,
KA Sims M., Smallwood N., Smith A., Smith M., Soutter M., Staden R.,
KA Watson J., Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K.,
KA Watson K., Watson A., Weinstock L., Wilkinson-Sproat J.,
FT "2.5 Mb of contiguous nucleotide sequence from chromosome III of C.
EL elegans.";
RL Nature 423:32-38(1994).
RN [2]
RP REVISIONS.
RA Durbin R.;
RI Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC
CC EMBL: Z27079; CAAL594.1;
DR PIR: FR8563; FR8563.
DR HSSP: P01132; IEPG.
DR WormPep: T05G5.1; CE21151.
DR InterPro: IPR006210; IEGF.
DR SMART: SM00181; EGF; 1.
KW Hypothetical protein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 412 HYPOTHETICAL 49.5 KDa PROTEIN T05G5.1.
SU SEQUENCE 412 AA; 49520 MW; 41A4D68C151918BA CPC64;

Query Match 1.28; Score 8; DB 1; Length 412;
Best Local Similarity 100.0%; Pred No R 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ILILLILMF 13
DB 5 ILILLILMF 12
IIIIIIII

RESULT 14
MER3_HUMAN
ID MER3_HUMAN STANDARD; PRT; 507 AA.
AC Q13064;
DT 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 15-SEP-2003 (Rel. 42, Last annotation update)
DE Makrin-3 (Zinc finger protein 127).
GN MKRN3 OR ZNF127.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE 93211018; PubMed=10196367;
PA Jung M.T.-C., Gray T.A., Ji Y., Glenn C.C., Saitoh S., Driscoll D.J.,
KA Nicholls R.D.;
RT "A novel imprinted gene, encoding a RING zinc finger protein, and
SC overlapping antisense transcript in the Prader-Willi syndrome
RL critical region.";
RL Hum. Mol. Genet. 8:783-793(1999).
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 3 CCH1-type zinc fingers.
CC

```

cc This SWISS prot entry is copyright. It is produced through a collaboration
cc between the Swiss Institute of Bioinformatics and the EMBL outstation.
cc The European Bioinformatics Institute. There are no restrictions on its
cc use by non-profit institutions as long as its content is in no way
cc modified and this statement is not removed. Usage by and for commercial
cc entities requires a license agreement (See <http://www.isi.ch/~bioinformatics/>
cc or send an email to license@isb.sib.ch).

DR EMBL: U13167; AAC13989.1;
DR PIR: J01614; J01614.
DR GenBank: J01614; MRPN1.
DR MIM: 608056;
DR Gene: 608056;
DR InterPro: IPR000571; Znf_CCH.
DR InterPro: IPR001841; Znf_CCH.
DR Pfam: PF00997; Zf-CCHC4.1.
DR Pfam: PF00642; Zf-CCHC3.1.
DR SMART: SM00184; PING: 1.
DR SMART: SM00356; Znf_CCH1.4.
DR PROSITE: PS00518; ZF_RING_1.1.
DR PROSITE: PS00896; ZF_RING_2.1.
KW Zinc finger, Repeat.
FT ZN FING 101 119 CCH1-TYPE 1.
FT ZN FING 244 262 CCH1-TYPE 2.
FT DOMAIN 266 293 MAKORIN-TYPE CYS HTS.
FT ZN FING 411 465 RING-TYPE.
FT ZN FING 469 420 CCH1-TYPE 3.
SQ SEQUENCE 507 AA; 55644 MW; 2E8FA91190F3A292 CRC64;

Query Match 1.2%; Score 8; DR 1; Length 507;
Best Local Similarity 100.0%; Prod. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 SEQUENCE 249
DB 11111111
608 SEQUENCE 375

RESULTS 14

ID SIB2 HUMAN STANDARD; PRT; 540 AA.
AC Q14425; Q14425.
DI 28 FEB 2003 (rel. 41, Created)
DI 28 FEB 2003 (rel. 41, Last sequence update)
DI 28 FEB 2003 (rel. 41, Last annotation update)
DE beta-2-syntrophin (5 kDa dystrophin-associated protein A1, basic
DE component 2) (Syntrophin 5) (SNT5) (Syntrophin-like) (SNTL).
GN SNTB2; SNT2B2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI TaxID 9606;
RN 11111111
RP HUMAN
RC TISSUE Muscle;
RX MEDLINE 94162617; PubMed 8576247;
RA Aho A.L., Feener C.A., Gussoni E., Yoshida M., Ozawa E., Kunkel L.M.;
RT "The three human syntrophin genes are expressed in diverse tissues,
RT have distinct chromosomal locations, and each bind to dystrophin and
RT its relatives.";
RL J. Biol. Chem. 271:2724-2730(1996).
RN 121
RP SEQUENCE FROM H.A. (13-FEB-02), AND INTERACTION WITH DMD; DTRA AND
RC TISSUE Brain;
RX MEDLINE 20496108; PubMed 11943403;
RA Ort J., Madsen-Olesen F., Borck B., Kuchinsky A.M., Borck S.;
RT "The receptor tyrosine phosphatase-like protein ICA512 binds the PDZ
RT domains of beta-2-syntrophin and ams in pancreatic beta cells.";
RL Eur. J. Cell Biol. 79:621-630(2000).
RN 131
RP INTERACTION WITH ERM64.

RX MEDLINE 20520281; PubMed 1075495;
RA Garcia R.A., Vasdevan K., Foundation A.;
RT "The neuroligin receptor ErbB 4 interacts with FcZ containing proteins
RT at neuronal synapses.";
RN Proc. Natl. Acad. Sci. U.S.A. 97:3366-3661(2000).
RN 141
RP INTERACTION WITH PIPKN.
RX MEDLINE 2176447; PubMed 11486065;
RA Ort J., Voronov S., Guo J., Zawalich K., Froehner S.C., Zawalich W.;
RA Solimena M.;
RT "Phosphorylation of beta2 syntrophin and Ca2+/Zn calpain mediated
RT cleavage of 17A12 upon stimulation of insulin secretion.";
RL EMBO J. 20:4013-4023(2001).
CC -1- FUNCTION: Adapter protein that binds to and probably organizes the
CC subcellular localization of a variety of membrane proteins. May
CC link various receptors to the actin cytoskeleton and the
CC dystrophin-glycoprotein complex. May play a role in the regulation
CC of secretory granules via its interaction with PIPKN.
CC -1- SUBUNIT: Monomer and homodimer (Probable). Interacts with the
CC other members of the syntrophin family: SNT1 and SNT4; and with
CC the sodium channel proteins SCN5A and SCN9A. Interacts with SASS1,
CC MAST205, microtubules and microtubule-associated proteins (by
CC similarity). Interacts with the dystrophin protein DMD and related
CC proteins DTRA and DTRN, and with the neuroligin receptor ERM64.
CC Interacts with PIPKN when phosphorylated, protect PIPKN from
CC protein cleavage by CAPN1. Phosphorylation upon insulin
CC stimulation disrupts the interaction with PIPKN and results in the
CC cleavage of PIPKN.
CC -1- SUBCELLULAR LOCATION: Membrane associated. In muscle, it is
CC exclusively localized at the neuromuscular junction (by
CC similarity). In insulinoma cell lines, it is enriched in secretory
CC granules.
CC -1- ALTERNATIVE PRODUCTS:
CC Frag. A1, read five aligned. Named isoforms 2
CC Name 1: Synonyms beta2 syntrophin-2;
CC IsoId: Q14425-1; Sequence Displayed:
CC Name-2: Synonyms beta2 syntrophin-2;
CC IsoId: Q14425-2; Sequence VSP: 006096, VSP: 006099;
CC Note: Lacks domains required for interaction with dystrophin-
CC related proteins;
CC 1- TISSUE SPECIFICITY: ubiquitous. Isoform 1 is the predominant
CC isoform. Weak level of isoform 2 is present in all tested tissues,
CC except in liver and heart where it is highly expressed.
CC -1- DOMAIN: The PD domain mediates the colocalization in a calcium
CC dependent manner (by similarity).
CC -1- DOMAIN: The FcZ domain binds to the last three of four amino acids
CC of ion channels and receptor proteins, the association with
CC dystrophin or related proteins probably leaves the FcZ domain
CC available to recruit proteins to the membrane (by similarity).
CC -1- DOMAIN: The SD domain binds calmodulin in a calcium dependent
CC manner (by similarity).
CC -1- PTM: Phosphorylated. Partially dephosphorylated upon insulin
CC stimulation.
CC 1- SIMILARITY: BELONGS TO THE SYNTROPHIN FAMILY.
CC 1- SIMILARITY: Contains 1 PDZ/ZIP domain.
CC 1- SIMILARITY: Contains 2 PD domains.
CC 1- SIMILARITY: Contains 1 SD (syntrophin unique) domain.
CC This SWISS-prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isi.ch/~bioinformatics/>
CC or send an email to license@isb.sib.ch).

DR EMBL: U40572; AAC10449.1;
DR EMBL: AF243485; AAK15147.1;
DR BSSD: P29476; 10AV.
DR GenBank: HGNC:11116; SNTB2.
DR MIM: 600027;
DR GO: 0016010; Dystrophin associated aligned at complex; IAS.
DR GO: 0005624; Membrane fraction; IAS.

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

[illegible][illegible]


```

DB 301 PARAPPDYDPTTETKGGGAGGAFESIVAFLEKIMAVTSEDEL 444
RESULT 6
Q81871
ID Q81871 PRELIMINARY: PRT: 660 AA.
AC Q81871:
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DI 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DI 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF 7 predictor
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024057; PubMed=1594074;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome."
RL Virology 185:120-131(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1594074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wano C.-K., Shikata T., Fukawa M., Fikihisa T., Mizuno K., Win K.M.;
RT "Hepatitis E virus: cDNA cloning and expression."
RL Microbiol. Immunol. 36:67-79(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93335008; PubMed=1636924;
RA Aye T.T., Uchida T., Ma X.Z., Iida F., Shikata T., Zhuang H.,
RA Win K.M.;
RT "Complete nucleic acid sequence of a hepatitis E virus isolated from the
RT Xinjiang epidemic (1986-1988) of China."
RL Nucleic Acids Res. 20:3512-3512(1992)
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.U., Reyes G.P., Tsareva T.S., Legters L.L.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus."
RL Proc Natl Acad Sci U S A 89:560-563(1992)
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92271462; PubMed=1580964;
RA Fry K.E., Tam A.W., Smith M.W., Kim J.P., Luk K.-C., Young L.M.,
RA Flatak M., Feldman R.A., Yun K.Y., Purdy M.A., McCausland K.A.,
RA Bradley D.W., Reyes G.R.;
RT "Hepatitis E virus (HEV): serial variation in the nonstructural gene
RT region encoding consensus motifs for an RNA-dependent RNA polymerase
RT and an ATP/GTP binding site."
RL Virus Genes 6:173-185(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=93348763; PubMed=8346669;
RA Bi S.-L., Purdy M.A., McCausland K.A., Margolis H.S., Bradley D.W.;
RT "The sequence of hepatitis E virus isolated directly from a single
RT source during an outbreak in China."
RL Virus Res. 29:233-247(1993).
DR EMBL: L08816; AAA04191.1;
DR InterPro: IPR004261; SP2; 1;
DR Pfam: PF03014; SP2; 1;
DR KEGG: K01101;
DR SIGNAL
FT SIGNAL 1 19 POTENTIAL.
SQ SEQUENCE 660 AA: 7699 MW: 65923.65 PII=614 Q8C64.
Query Match 43.0% Score 284, LE 12, Length 660;
Best Local Similarity 100.0%; Pred. No. 2,66-266;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

```

```

245 VVNSVRLDVAQVQWRSVLSNVAFLEAAGSLVWVLRVLAQVSLRIVLTAAGLLELAL 204
|||||
246 VVPSLEPLRYENCWPSVETSVVAEEEAATSCVWVLRVPSVNSVYNTVYFATGILDEAL 404
|||||
247 ELLEPSLZTRITRIVRSSTETARHELEKACAELETTAAATTRKDLATFISNGVREI 404
|||||
248 ELLEPRLTVGNTNTRVRSYSSSTARHRIKGAAGTAAELTTAATSPMKDLYELFISNGVREI 404
|||||
249 GEGATLTENLATLGLSLPTLSSAALLEYSFVVSANSTLVETTFVFNKQLGK 424
|||||
250 GEGATLTENLATLGLSLPTLSSAALLEYSFVVSANSTLVETTFVFNKQLGK 424
|||||
251 IALPHDINDGESPVQIQWQDNEHQDRTPSPAPSPFSVLKANIWI SLTAAEYQAST 484
|||||
252 IALPHDINDGESPVQIQWQDNEHQDRTPSPAPSPFSVLKANIWI SLTAAEYQAST 484
|||||
253 YASSTGTPVVVSQSVTLNVNATGAQAVAPSLWTKVILDERPLST 528
|||||
254 YASSTGTPVVVSQSVTLNVNATGAQAVAPSLWTKVILDERPLST 528
|||||
RESULT 7
Q9WLL4
ID Q9WLL4 PRELIMINARY: PRT: 660 AA.
AC Q9WLL4:
DI 01-NOV-1999 (TrEMBLrel. 12, Created)
DI 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DI 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93175752;
RA Yeh C.V., Fuchting M., Packer M.J., Leber C.P., Davis R.L.;
RT "Hepatitis E virus: complete genome sequence and phylogenetic analysis
RT of a Nepali isolate."
RL Virus Res. 57:21-26(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX EPIFAM T15792;
RA Yeh C.V.;
RT Submitted (MAR-1998) to the EPIFAM database, type: data bases.
DR EMBL: AF051830; AAC97188.1;
DR Fasta: AF051830; SP2; 1;
DR Pfam: PF03014; SP2; 1;
SQ SEQUENCE 660 AA: 76798 MW: 60400.85A33085LE Q9C64;
Query Match 43.7% Score 292, LE 12, Length 660;
Best Local Similarity 100.0%; Pred. No. 2,36-264;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

```

```

235 VNYTHTTTCAGGGLHALLDEERLFGGRHVSRYSTSLARHLPKGAIGAEIIT 444
|||||
236 VNSYTNPTVTCAGLGLLEALEPEPNTNPVSPYSSTARHPIPKGAIGAEIIT 444
|||||
237 AATPEWELYSSTNVYFICHTALFPAALGLLGLPTLSSAALYFSRPVVA 404
|||||
238 AATPEWELYSSTNVYFICHTALFPAALGLLGLPTLSSAALYFSRPVVA 404
|||||
239 RPTPTVYVTVNATLWPKATLPHPTIGPFWVYVWQHQDPSPAPSPFSV 464
|||||
240 NPEPTVKVTSVENAQKGIAPRHHGSPKVVVQVWQHQDPSPAPSPFSV 464
|||||
241 LKAGVRLKELANATLWQSTELSTNTVVSANVTVWALVAVASLLWVILDER 524
|||||
242 LKANIIVLWLSIAELAQSSAGVVSAGVVSAGVVSAGVVSAGVVSAGVVS 524
|||||
243 LKPTVTLVETFTVTLVETFTVTLVETFTVTLVETFTVTLVETFTVTLV 524
|||||

```



```
Matches 559; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 99 RRRRTTAAAFELTAVAFADHTPPVPVDSRGAILPRQYNISSPLISSVATGTLNLVYAA 158
DB 99 RRRRTTAAAFELTAVAFADHTPPVPVDSRGAILPRQYNISSPLISSVATGTLNLVYAA 158
QY 159 PLSPLLPLQDCTNHIMATEASNAQYRVARATIRYKPLVPNAVGGYALISISFWPQTIT 218
DB 159 PLSPLLPLQDCTNHIMATEASNAQYRVARATIRYKPLVPNAVGGYALISISFWPQTIT 218
QY 219 PTVSMNSITSTVPIIVQPGIASELVIPSERLHPNQGWRSVETSGVAEEEAISGLVNL 278
DB 219 PTVSMNSITSTVPIIVQPGIASELVIPSERLHPNQGWRSVETSGVAEEEAISGLVNL 278
QY 279 CIHGLSVNSYNTPTGALGLDFALELEPRLNTPGNTNTRVSPYSSSTAPRHLPRCADGT 338
DB 279 CIHGLSVNSYNTPTGALGLDFALELEPRLNTPGNTNTRVSPYSSSTAPRHLPRCADGT 338
QY 339 AELTTTAAATPMKPLVFTSNVSGEIGPGIALTLFNLAVTLGLGPTLTISSAGGOLFYS 398
DB 339 AELTTTAAATPMKPLVFTSNVSGEIGPGIALTLFNLAVTLGLGPTLTISSAGGOLFYS 398
QY 399 RPVVSANGCEPTVKLYTSVENAQQDKGLAIPHIDILGESKVVTDYDQHQHQKPTPSAP 458
DB 399 RPVVSANGCEPTVKLYTSVENAQQDKGLAIPHIDILGESKVVTDYDQHQHQKPTPSAP 458
QY 459 SRPESVLRANDVLMWLSLTAAYDQSTYSGSIGPVVYSDSVTLNVNVAIGACAVARSLDWIK 518
DB 459 SRPESVLRANDVLMWLSLTAAYDQSTYSGSIGPVVYSDSVTLNVNVAIGACAVARSLDWIK 518
QY 519 VTLGRPLSTLOQYSKTFEVLPLRGLKSEWAGTTKAGYPYNYNTASDQLLVNNAAGHR 578
DB 519 VTLGRPLSTLOQYSKTFEVLPLRGLKSEWAGTTKAGYPYNYNTASDQLLVNNAAGHR 578
QY 579 VAISTYTTSLGAGPVSISAVAVLAPHSAALALFETIETVPAPAHIDFCEPCEPPLGQGC 638
DB 579 VAISTYTTSLGAGPVSISAVAVLAPHSAALALFETIETVPAPAHIDFCEPCEPPLGQGC 638
QY 639 AFQSTVAELQRLKMKVGKTRLEL 660
DB 639 AFQSTVAELQRLKMKVGKTRLEL 660

RESULT 11
ID Q91855 PRELIMINARY: PRT: 660 AA.
AC Q91855:
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN:GCM.
RA MEDLINE:990105430; PubMed:6890424;
RA Isaac S.A., Binn L.N., Gumatesi P.J., Arthur R.K., Meade M.K.,
RA van Geyck-Gandre H., Longor G.F., Innis B.L.;
PT "Phylogenetic analysis of hepatitis E virus isolates from Egypt.";
RL J. Med. Virol. 57:68-74(1999).
DR EMBL: AF051351; AAC35761.1; -.
DR InterPro: IPR004261; SP2;
DR Pfam: PF03014; SP2; 1.
SQ SEQUENCE 660 AA: 71040 MW: 64485A76492791 CpGc1;

Query Match: 39.0%; Score 251; DB 12; Length 660;
Best Local Similarity 99.5%; Pred. No. 2,78-234;
Matches 561; Conservative 0; Mismatches 3; Indels 0; Gaps 0.
QY 11 LMFLPMLPAPPQPSPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 70
```

```
DB 11 LMFLPMLPAPPQPSPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 70
QY 71 GAGPRVQAPAPLASSAMFQAGSPAVASRRRTTAAAGLTAFAFAHDTFTVFWVWSSIA 130
DB 71 GAGPRVQAPAPLASSAMFQAGSPAVASRRRTTAAAGLTAFAFAHDTFTVFWVWSSIA 130
QY 71 GAGPRVQAPAPLASSAMFQAGSPAVASRRRTTAAAGLTAFAFAHDTFTVFWVWSSIA 130
DB 71 GAGPRVQAPAPLASSAMFQAGSPAVASRRRTTAAAGLTAFAFAHDTFTVFWVWSSIA 130
QY 131 LLKRONLSPLSTSSVATCTNLVLYAAPLSPLPLQKTNTHIMATEASNAQYRVARAKA 190
DB 131 LLKRONLSPLSTSSVATCTNLVLYAAPLSPLPLQKTNTHIMATEASNAQYRVARAKA 190
QY 131 LLKRONLSPLSTSSVATCTNLVLYAAPLSPLPLQKTNTHIMATEASNAQYRVARAKA 190
DB 131 LLKRONLSPLSTSSVATCTNLVLYAAPLSPLPLQKTNTHIMATEASNAQYRVARAKA 190
QY 191 TIRYKPLVPNAVGGYALISISFWPQTITPTISVDMNSITSTVPIIVQPGIASELVIPSE 250
DB 191 TIRYKPLVPNAVGGYALISISFWPQTITPTISVDMNSITSTVPIIVQPGIASELVIPSE 250
QY 191 TIRYKPLVPNAVGGYALISISFWPQTITPTISVDMNSITSTVPIIVQPGIASELVIPSE 250
DB 191 TIRYKPLVPNAVGGYALISISFWPQTITPTISVDMNSITSTVPIIVQPGIASELVIPSE 250
QY 251 LHYRNOGWRSVETSGVAEEEAISGLVNLGCHSLVNSYNTPTGALGLDFALELEPRLN 310
DB 251 LHYRNOGWRSVETSGVAEEEAISGLVNLGCHSLVNSYNTPTGALGLDFALELEPRLN 310
QY 311 LTPONTNTRVSKYSSTARRHLRGAAGTAAELTTTAAATPMKPLVFTSNVSGEIGP 370
DB 311 LTPONTNTRVSKYSSTARRHLRGAAGTAAELTTTAAATPMKPLVFTSNVSGEIGP 370
QY 371 TIENADTLGLRPTLTISSAGGLPYSRPVVSANGTPTVKLYTSVENAQQDKGLAIPH 430
DB 371 TIENADTLGLRPTLTISSAGGLPYSRPVVSANGTPTVKLYTSVENAQQDKGLAIPH 430
QY 371 TIENADTLGLRPTLTISSAGGLPYSRPVVSANGTPTVKLYTSVENAQQDKGLAIPH 430
DB 371 TIENADTLGLRPTLTISSAGGLPYSRPVVSANGTPTVKLYTSVENAQQDKGLAIPH 430
QY 431 IDGESRVVITQPVNDHQEDRPTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 490
DB 431 IDGESRVVITQPVNDHQEDRPTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 490
QY 491 PVVVSNSVTLVNVATCAAVARSLDWIKVYKVTGDPPISTIDQYKSPKEVPIPRKLS 550
DB 491 PVVVSNSVTLVNVATCAAVARSLDWIKVYKVTGDPPISTIDQYKSPKEVPIPRKLS 550
QY 551 GTTKAGYPYNYNT 564
DB 551 GTTKAGYPYNYNT 564

RESULT 12
QYWM10 PRELIMINARY: PRT: 605 AA.
AC QYWM10:
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Structural protein (Fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN:HCN.
RA MEDLINE:99013657; PubMed:9797311;
RA Pina S., Joffre J., Emerson S.O., Purcell R.H., Gitones R.;
PT "Characterization of a strain of infectious hepatitis E virus isolated
PT from sewage in an area where hepatitis E is not endemic.";
RL Aspl. Infect. Microbiol. 64:4485-4488(1998).
DR EMBL: AF059884; AAC77408.1; -.
DR InterPro: IPR004261; SP2;
DR Pfam: PF03014; SP2; 1.
DR NON_TER 605 605
SQ SEQUENCE 605 AA: 74740 MW: A08017197182972 CR664;

Query Match: 97.4%; Score 237; DB 12; Length 605;
Best Local Similarity 100.0%; Pred. No. 1,98-240;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY 99 RNVTTSEHATLPPNATLTISSPLISSVATGTLNLVYAA 418
DB 99 RNVTTSEHATLPPNATLTISSPLISSVATGTLNLVYAA 418
QY 99 RNVTTSEHATLPPNATLTISSPLISSVATGTLNLVYAA 418
DB 99 RNVTTSEHATLPPNATLTISSPLISSVATGTLNLVYAA 418
```


Matches	197;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	436	DGTAELTTAATREMKDIYFTSTNGVGFGRGTATLFLNLAOTLLGSLPTELISSAGGDL	395						
Db									
QY	446	DGTAELTTAATREMKDLYFTSINGVGEIGRGIALTLFLNLADTLGLGLPTELISSAGGDL	395						
Db									
QY	496	FYSRPVVSANGPEPIVKLYTSVENAQCKGIALPHDIDLGESRVVLDYDQNHQHQDRPTPS	455						
Db									
QY	496	FYSRPVVSANGPEPIVKLYTSVENAQCKGIALPHDIDLGESRVVLDYDQNHQHQDRPTPS	455						
Db									
QY	456	PAPSRPFSLRANVWLSLTAAEYDQSTYGSSTGPVYVSDSVILYNVAIGACAVARSLD	515						
Db									
QY	456	PAPSRPFSLRANVWLSLTAAEYDQSTYGSSTGPVYVSDSVILYNVAIGACAVARSLD	515						
Db									
QY	516	WTKVTLGDRPLSTIQY	532						
Db									
QY	516	WTKVTLGDRPLSTIQY	532						
Db									

Search completed: August 5, 2003, 09:33:41
Job time : 48 secs


```

XX (GENE-) GENELABS TECHNOLOGIES INC.
PA (USSH ) US SEC DRPT HEALTH.
XX Bradley DW, Carl M, Reyes GR, Tam AW;
XX
XX WPI: 1993-243144/30.
DR N-PSDR: AAQ47129
XX
XX New immunogenic hepatitis E virus (HEV) peptides are selected from the
PI ORF1, ORF2 and ORF3 regions of HEV. The peptides can be used in
PI vaccines to prevent infection by HEV. The antibodies can neutralise
PI and block HEV infection and can be used to prevent or treat HEV
PI infection. The peptides and antibodies can also be used as
XX diagnostic reagents.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 660 AA:
Query Match 100.0%; Score 3408; DB 14; Length 660;
Best Local Similarity 100.0%; Pred. No. 8,60-263;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 1 MPPPTILLILMLFPMIPAPPNGQSGTFRPPSSGSGGTPWQWISGPPAIPYHPTN 60
DB 1 MPPPTILLILMLFPMIPAPPNGQSGTFRPPSSGSGGTPWQWISGPPAIPYHPTN 60
QY 61 PPAPVTAACAGCPVQPPAPPGISAWPGCAQHPAVASPPPTTACAAPLIAVAHAHTP 120
DB 61 PPAPVTAACAGCPVQPPAPPGISAWPGCAQHPAVASPPPTTACAAPLIAVAHAHTP 120
QY 121 PVPDVDSRGATLRQYVNLSTSPITSSVATGTLNVLVAAPLSPLPLQDGTNTHIMATEAS 180
DB 121 PVPDVDSRGATLRQYVNLSTSPITSSVATGTLNVLVAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQYVAPATIRYPLPVNAVSGYATISFPWQTTTTTSDVMNSITSTVPEILVQGTI 240
DB 181 NYAQYVAPATIRYPLPVNAVSGYATISFPWQTTTTTSDVMNSITSTVPEILVQGTI 240
QY 241 ASELVIPSERLHYPNQWPSVETSGVAEEEAATSGLVMLCHGSLVNSYNTPTVGAIGLL 300
DB 241 ASELVIPSERLHYPNQWPSVETSGVAEEEAATSGLVMLCHGSLVNSYNTPTVGAIGLL 300
QY 301 DFALELEFERNITPGNTNTRVSPYSSTAPRPLRPGAGTAELTTTAATPRPMKMLYPTING 360
DB 301 DFALELEFERNITPGNTNTRVSPYSSTAPRPLRPGAGTAELTTTAATPRPMKMLYPTING 360
QY 361 VSEIGRGIALFLNLIADTLIGGLPTFLISSAGQQLYFSPVVSANCFPTVKLYTSVENAQ 420
DB 361 VSEIGRGIALFLNLIADTLIGGLPTFLISSAGQQLYFSPVVSANCFPTVKLYTSVENAQ 420
QY 421 GKKGIAIPIHDIGESVWYIQYVQHQEPTPTPSAPSPSPSVIPANAVIMLSITAAEY 480
DB 421 GKKGIAIPIHDIGESVWYIQYVQHQEPTPTPSAPSPSPSVIPANAVIMLSITAAEY 480
QY 481 DOSTYGSSTGFVVSQSVILVNVATCAQAVASLWTKVTLQDPLSLSLGYSPTFVFI 540
DB 481 DOSTYGSSTGFVVSQSVILVNVATCAQAVASLWTKVTLQDPLSLSLGYSPTFVFI 540
QY 541 IFFPLSTWEACTFKAYFYWYNTATGQIIVENAAHPVAISTYITSTGATVSSAVAV 600
DB 541 IFFPLSTWEACTFKAYFYWYNTATGQIIVENAAHPVAISTYITSTGATVSSAVAV 600
QY 601 IAPHCALIPETIENAPAHITFCPTPTPTFCIGQAFQSTVAELPLAKMKVSKLEL 660
DB 601 IAPHCALIPETIENAPAHITFCPTPTPTFCIGQAFQSTVAELPLAKMKVSKLEL 660

```

RESULT 2

```

AAR39406
XX AAR39306 standard; Protein; 660 AA.
XX
XX AAR39306;
XX
XX 25 MAR-2003 (updated)
DT 14-FEB-1994 (first entry)
XX
XX Burma strain HEV ORF2 putative virus capsid protein.
XX
XX Hepatitis E virus; vaccine; neutralising antibodies; infection;
XX block; open reading frame; antibodies.
XX
XX Hepatitis E virus.
XX
XX WPI:1993-243227/30.
XX 22-JUL-1994.
XX
XX 19-JAN-1992; 93W: US200175.
XX
XX 17-JAN-1992; 92US-0822435.
XX 20-APR-1992; 92US-0870985.
XX
XX (GENE ) GENELABS TECHNOLOGIES INC.
XX (USSH ) US DRPT HEALTH & HUMAN SERVICES.
XX
XX Bradley DW, Elisei J, Esaki K, Field M, Reyes GR, Tam AW, Tam J;
XX
XX WPI: 1993-243227/30.
XX N-PSDR: AAQ46813.
XX
XX Antigen and antibody vaccines against hepatitis E virus infection
XX - contain peptides derived from capsid protein C-terminal or
XX antibodies against protein
XX
XX Disclosure; Fig 7; 43pp; English.
XX
XX The sequence is that of the putative virus capsid protein encoded
XX by Burma strain hepatitis E virus (HEV) open reading frame ORF2.
XX This protein or peptide fragments of it may be used in a vaccine
XX composition for immunising an individual against HEV. Antibodies
XX raised against these peptides can also be used in such vaccines.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 660 AA:
Query Match 100.0%; Score 3408; DB 14; Length 660;
Best Local Similarity 100.0%; Pred. No. 8,60-263;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 1 MPPPTILLILMLFPMIPAPPNGQSGTFRPPSSGSGGTPWQWISGPPAIPYHPTN 60
DB 1 MPPPTILLILMLFPMIPAPPNGQSGTFRPPSSGSGGTPWQWISGPPAIPYHPTN 60
QY 61 PPAPVTAACAGCPVQPPAPPGISAWPGCAQHPAVASPPPTTACAAPLIAVAHAHTP 120
DB 61 PPAPVTAACAGCPVQPPAPPGISAWPGCAQHPAVASPPPTTACAAPLIAVAHAHTP 120
QY 121 PVPDVDSRGATLRQYVNLSTSPITSSVATGTLNVLVAAPLSPLPLQDGTNTHIMATEAS 180
DB 121 PVPDVDSRGATLRQYVNLSTSPITSSVATGTLNVLVAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQYVAPATIRYPLPVNAVSGYATISFPWQTTTTTSDVMNSITSTVPEILVQGTI 240
DB 181 NYAQYVAPATIRYPLPVNAVSGYATISFPWQTTTTTSDVMNSITSTVPEILVQGTI 240
QY 241 ASELVIPSERLHYPNQWPSVETSGVAEEEAATSGLVMLCHGSLVNSYNTPTVGAIGLL 300
DB 241 ASELVIPSERLHYPNQWPSVETSGVAEEEAATSGLVMLCHGSLVNSYNTPTVGAIGLL 300

```


Reyes GR, Yarbough PG, Bradley DW, Krawczynski KZ, Tam AW, Fry KE;
 WPL: 2001; 44:705/56.
 N PSUB: AAP94493.
 New DNA sequences of enterically transmitted non-A/non-B (ET-NANB) hepatitis viral agent, useful in diagnosing infection by an enterically transmitted agent (e.g. ET-NANB virus), as well as in vaccine production.
 Disclosure: Columns 64-68; 45pp; English.
 The invention relates to an isolated DNA comprising the genome of an enterically transmitted non-A/non-B (ET-NANB) viral hepatitis agent (also referred as HEV). The DNA sequences or their fragments are useful in preparing ET-NANB proteins and as probes for virus detection. These are particularly useful in diagnosing infection by an enterically transmitted agent (e.g. ET-NANB virus), as well as in vaccine production. The present sequence represents a ORF2 protein encoded by a ET-NANB viral DNA sequence from HEV-Burma strain.
 Sequence: 660 AA;

20017 M¹-ch 100.0% Score: 4406; DB:22; Length: 660;
 Best Local Similarity 100.0%; Pred. No. gaps: 263;
 Matches: 660; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

1 MKPRFTLLLEMLFAPPPGSGSRGRKSGSGGFWGDRVDSQPPALTYDPTN 60
 1 MKPRFTLLLEMLFAPPPGSGSRGRKSGSGGFWGDRVDSQPPALTYDPTN 60
 41 PRAPVATAAGAGPVPGFARPLGSAWPDQAGRIAVASRRPPTTAGAAPLTAVAPADTP 120
 41 PRAPVATAAGAGPVPGFARPLGSAWPDQAGRIAVASRRPPTTAGAAPLTAVAPADTP 120
 121 PWGDSKALLRRGYNSISLSSVAATGGINVLYAAPLSPLPLQGVNTHIMALEAS 180
 121 PWGDSKALLRRGYNSISLSSVAATGGINVLYAAPLSPLPLQGVNTHIMALEAS 180
 181 NYAGYVAPATIRYRPLVNAVGGYALISFWPQTTTTPTTSVMNSITSTAVRLVPGGI 240
 181 NYAGYVAPATIRYRPLVNAVGGYALISFWPQTTTTPTTSVMNSITSTAVRLVPGGI 240
 241 ASLLVTPSRHYRNGWRSVETSVARFEATSGLMVCHGSLVNSYNTPTPYGALGL 300
 241 ASLLVTPSRHYRNGWRSVETSGVAEEATSGLMVCHGSLVNSYNTPTPYGALGL 300
 301 DEALELEFNLTPTGNTIRVSKYSTAHRLRPGATGTAELTTTAATPFMKDLYFTSTNG 360
 301 DEALELEFNLTPTGNTIRVSKYSTAHRLRPGATGTAELTTTAATPFMKDLYFTSTNG 360
 361 VPIFGGIALTFNLADTLGGTLPELSSAGGGLFYSGPVVVSANRPTTKRIYTSVENAQ 420
 361 VPIFGGIALTFNLADTLGGTLPELSSAGGGLFYSGPVVVSANRPTTKRIYTSVENAQ 420
 421 GAKGIALPHDGLHGSPPVVLQVLDNREKQKPTLSPALSPFESVLRANDVLSLAAEY 480
 421 GAKGIALPHDGLHGSPPVVLQVLDNREKQKPTLSPALSPFESVLRANDVLSLAAEY 480
 481 DQSTYGSSITGPVYVSUSVILVNVAIGAAVAKSTLTKVTLDPPTLSTAGYSKTFPVLD 540
 481 DQSTYGSSITGPVYVSUSVILVNVAIGAAVAKSTLTKVTLDPPTLSTAGYSKTFPVLD 540
 541 LRKGLSPWATTKACVYNNNTASIGLLVFNAGCHEVALISTYTTSLGAGPVSISAVV 600
 541 LRKGLSPWATTKACVYNNNTASIGLLVFNAGCHEVALISTYTTSLGAGPVSISAVV 600
 601 LAHSALALLELTLYPAPARHTFQPTKPLGLQYKAFQSTVARELQPLKMKVSKTPEL 660
 601 LAHSALALLELTLYPAPARHTFQPTKPLGLQYKAFQSTVARELQPLKMKVSKTPEL 660

RES01.1 9

AA015699
 ID AA015699 standard; Protein; 660 AA.
 XX AA
 AA AA015699;
 XX AA
 09-NEW-2002 (first entry)
 XX
 Hepatitis E virus (Burma strain) ORF2 protein.
 XX
 HEV; enterically transmitted nonA/nonB hepatitis viral agent; Burma strain; file: ORF2.
 XX
 Hepatitis E virus.
 XX
 US6479891-B1.
 XX
 40-APR-2002.
 XX
 19-APR-2000; 2000US 0554427.
 XX
 25-JUN-1994; 940US 0279824.
 XX
 07-JUN-1995; 950US 0474507.
 XX
 15-APR-1991; 910US 0661078.
 XX
 17-JUN-1988; 880US 0208947.
 XX
 11-APR-1989; 890US 0346672.
 XX
 16-JUN-1989; 890US 0367486.
 XX
 13-DEC-1989; 890US 0420921.
 XX
 05-APR-1990; 900US 0505888.
 XX
 (ORF2) US PAT. HEALTH & HUMAN SERVICES.
 XX
 (ORF2) GENE:HEV:ORF2:ORF2:IN1.
 XX
 Reyes GR, Yarbough PG, Bradley DW, Krawczynski KZ, Tam AW, Fry KE;
 WPL: 2001; 44:705/56.
 N PSUB: AAP94493.
 Detecting the presence of enterically transmitted, nonA/nonB hepatitis viral (HEV) agents in bile samples from infected humans and monkeys using polymerase chain reaction
 XX
 Disclosure: Column 19-24; 61pp; English.
 XX
 The invention comprises a method for detecting the presence of enterically transmitted nonA/nonB hepatitis viral (HEV) agents in a sample and isolating HEV agents or nucleic acid fragments produced by the agent. The method utilizes PCR using bile from a human or cynomolgus monkey actively infected with HEV as a source of the agent. The method of the invention is used for detecting the presence of a viral agent in a sample of cultured cells infected with the agent and isolation of enterically transmitted nonA/nonB HEV agents or nucleic acid fragments produced by the agent. The present amino acid sequence represents the protein encoded by open reading frame 2 (ORF2) of a DNA sequence isolated from a Burmese strain of the hepatitis E virus.
 XX
 Sequence: 660 AA;

Query Match Indels Score 4406; DB:22; Length: 660;
 Best Local Similarity 100.0%; Pred. No. gaps: 263;
 Matches: 660; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

QY 1 MKPRFTLLLEMLFAPPPGSGSRGRKSGSGGFWGDRVDSQPPALTYDPTN 60
 1 MKPRFTLLLEMLFAPPPGSGSRGRKSGSGGFWGDRVDSQPPALTYDPTN 60
 41 PRAPVATAAGAGPVPGFARPLGSAWPDQAGRIAVASRRPPTTAGAAPLTAVAPADTP 120
 41 PRAPVATAAGAGPVPGFARPLGSAWPDQAGRIAVASRRPPTTAGAAPLTAVAPADTP 120
 121 PWGDSKALLRRGYNSISLSSVAATGGINVLYAAPLSPLPLQGVNTHIMALEAS 180
 121 PWGDSKALLRRGYNSISLSSVAATGGINVLYAAPLSPLPLQGVNTHIMALEAS 180


```

121 PVPVDSGATIPQYNLSPLTSSVATGNTLVLAAPLSPLLQDGTNTHMATEAS 190
141 NYAQVAPATIPYPPVLPVNAVGYATISSEWPCPTPTTPTSVMMNSITSNWPTTVPQI 240
161 NYAQVAPATIPYPPVLPVNAVGYATISSEWPCPTPTTPTSVMMNSITSTDEILVQPI 240
241 ASFLVPSERLHYNQWBSVETSGVAFFATSCI VMIOTHSIVNSYNTIYTGALL 300
241 ASFLVPSERLHYNQWBSVETSGVAFFATSCI VMIOTHSIVNSYNTIYTGALL 300
301 DFALFEPRLTPGNTNTPVSPYSSTAPRPLRPGADGTAFITTAATPEMKDLYPTING 360
361 VGEIGRGIALTFINADITLGGIPTEPLISSAGGCLFYSRPVVSANGPTVKIYTSVENAQ 420
361 VGEIGRGIALTFINADITLGGIPTEPLISSAGGCLFYSRPVVSANGPTVKIYTSVENAQ 420
421 QDKGIATPHDIDGESKVVLIQYQUNQHEQKPTSPAPSPPSVLPVANDVWLSLTAABY 480
421 QDKGIATPHDIDGESKVVLIQYQUNQHEQKPTSPAPSPPSVLPVANDVWLSLTAABY 480
481 DOSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDWTKVTLQGRPLSTTQQYSKTFVLP 540
481 DOSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDWTKVTLQGRPLSTTQQYSKTFVLP 540
541 LRGLSPWEAGTTKAGYPYNYNTTASDOLLVENAAGHRVAISPTYTSLGAGPVSSAVAV 600
541 LRGLSPWEAGTTKAGYPYNYNTTASDOLLVENAAGHRVAISPTYTSLGAGPVSSAVAV 600
601 LAPHSALEDDTLDYPARARTDDPCPCRPPIGLAGCAPQSTVAELQPLKKVKVKTREL 660
601 LAPHSALEDDTLDYPARARTDDPCPCRPPIGLAGCAPQSTVAELQPLKKVKVKTREL 660

```

RESULT 13

AA91814
ID AA91814 standard: Protein, 660 AA.

AC AA91814:

DT 26-NOV-1996 (first entry)

XX Hepatitis E virus strain SAP-55 ORF 2.

XX Hepatitis E virus; HFV, SAP-55 strain; enteric transmission;

KW structural region; antigen; detection; antibody; vaccine;

KW immunisation; infection.

XX Hepatitis E virus

XX WO9610580-A2.

XX 11-APR-1996

XX 03-OCT-1995: 95WO-US13102.

XX 03-OCT-1994: 94US-0316765.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Emerson SU, Purcell RH, Tsarev SA.

XX WPI: 1996-209320/21.

XX N-ESDB: AAT27394.

XX Isolated and purified hepatitis E virus strain SAP-55 DNA - encodes
PT antigenic protein useful in diagnosis, prophylaxis and treatment of
PI hepatitis E virus infection

XX Claim 6: Pages 13-15; 121pp; English.

XX

CC The present sequence is the protein prod. of ORF 2 from the
CC hepatitis E virus (HEV) strain SAP-55, which was implicated in an
CC enterically transmitted non-A, non-B hepatitis in Pakistan. The
CC protein encoded by the structural region of the virus (i.e. ORF 2),
CC which is capable of forming HEV like particles, is useful for the
CC detection of HEV antibodies (pref. IgG or IgM) in blood, plasma,
CC sera, cerebrospinal fluid, tissue, urine or pleural fluid. The
CC protein, and anti-HEV antibodies generated using the protein, can
CC also be used in vaccines for immunising an animal against HEV
CC infection. The protein is identified as a band of greater than
CC 50 kDa following SDS PAGE of cell lysates of insect cells infected
CC with a HEV ORF 2 cont. baculovirus. i.e. the claimed recombinant
CC expression vectors pPIC9-1779, -1780 and -1781.

XX Sequence 660 AA:

Query Match 99.5% Score 3391; DB 17; Length 660;
Best Local Similarity 99.4%; Pct. Id. 26.261;
Matches 656; Conservative 1; Mismatches 3; Indels 0; gaps 0;

```

QY 1 MRDRPILLILLMLPMLPAPPQGPSRRRGKRSQSGSGPFWQVUSLPPFAIPYHPIN 60
DB 1 MRDRPILLILLMLPMLPAPPQGPSRRRGKRSQSGSGPFWQVUSLPPFAIPYHPIN 60
QY 61 PFAPDVTAAAGAPRVPRQAPRLSSAWRQQAAPPVASSRPPTTAAAPLTAAPAHPTP 120
DB 61 PFAPDVTAAAGAPRVPRQAPRLSSAWRQQAAPPVASSRPPTTAAAPLTAAPAHPTP 120
QY 121 PVDVVISSEGLIDRQVRLSSLLSSVATGTRNLVYAAPLSLLIQQGNTTHMATEAS 180
DB 121 PVDVVISSEGLIDRQVRLSSLLSSVATGTRNLVYAAPLSLLIQQGNTTHMATEAS 180
QY 181 NYAGYKVARATIDYRPIVNAVGYATISSEWPCPTPTTPTSVMMNSITSNWPTTVPQI 240
DB 181 NYAGYKVARATIDYRPIVNAVGYATISSEWPCPTPTTPTSVMMNSITSNWPTTVPQI 240
QY 241 ASFLVPSERLHYNQWBSVETSGVAFFATSCI VMIOTHSIVNSYNTIYTGALL 300
DB 241 ASFLVPSERLHYNQWBSVETSGVAFFATSCI VMIOTHSIVNSYNTIYTGALL 300
QY 301 DFALFEPRLTPGNTNTPVSPYSSTAPRPLRPGADGTAFITTAATPEMKDLYPTING 360
DB 301 DFALFEPRLTPGNTNTPVSPYSSTAPRPLRPGADGTAFITTAATPEMKDLYPTING 360
QY 361 VGEIGRGIALTFINADITLGGIPTEPLISSAGGCLFYSRPVVSANGPTVKIYTSVENAQ 420
DB 361 VGEIGRGIALTFINADITLGGIPTEPLISSAGGCLFYSRPVVSANGPTVKIYTSVENAQ 420
QY 421 QDKGIATPHDIDGESKVVLIQYQUNQHEQKPTSPAPSPPSVLPVANDVWLSLTAABY 480
DB 421 QDKGIATPHDIDGESKVVLIQYQUNQHEQKPTSPAPSPPSVLPVANDVWLSLTAABY 480
QY 481 DOSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDWTKVTLQGRPLSTTQQYSKTFVLP 540
DB 481 DOSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDWTKVTLQGRPLSTTQQYSKTFVLP 540
QY 541 LRGLSPWEAGTTKAGYPYNYNTTASDOLLVENAAGHRVAISPTYTSLGAGPVSSAVAV 600
DB 541 LRGLSPWEAGTTKAGYPYNYNTTASDOLLVENAAGHRVAISPTYTSLGAGPVSSAVAV 600
QY 601 LAPHSALEDDTLDYPARARTDDPCPCRPPIGLAGCAPQSTVAELQPLKKVKVKTREL 660
DB 601 LAPHSALEDDTLDYPARARTDDPCPCRPPIGLAGCAPQSTVAELQPLKKVKVKTREL 660

```

RESULT 14

AA91814
ID AA91814 standard: Protein, 660 AA.

XX AA91814:

XX 11-JUN-1999 (first entry)

XX

Db 61 PFAPDVTAAAGAGPVVRQAPPIGSAMWQAGPPAAASPPPTTAAAGAPLTAVAPAHDP 120
QY 121 PVPDYDSKAILRQYNLSTSPITSSVATGNTNLVLAAPLSPLPLQUGHNTNTHIMATEAS 180
Db 121 PVPDYDSKAILRQYNLSTSPITSSVATGNTNLVLAAPLSPLPLQUGHNTNTHIMATEAS 180
QY 181 NYAAYRVAPATIRPYEPVNAVGGYALISFWQTTTTPTSVFMNSITSDVPIILVQPGI 240
Db 181 NYAAYRVAPATIRPYEPVNAVGGYALISFWQTTTTPTSVFMNSITSDVPIILVQPGI 240
QY 241 ASELVIPSERIHYRNQWRSVETSGVAEREATSGLVMLCHGSLVNSYNTPTVTGALGLL 300
Db 241 ASELVIPSERIHYRNQWRSVETSGVAEREATSGLVMLCHGSLVNSYNTPTVTGALGLL 300
QY 301 DFALFPPNLTPOGNTNTRVSPYSSTAPRBRACAGTAEITTTAATREMKDLYFTSTNG 360
Db 301 DFALFPPNLTPOGNTNTRVSPYSSTAPRBRACAGTAEITTTAATREMKDLYFTSTNG 360
QY 361 VGEIGRGIAITFNIAADTILAHPTETISSAGGQLEFYSPEVVSANGPPTVKLYTSVENAQ 420
Db 361 VGEIGRGIAITFNIAADTILAHPTETISSAGGQLEFYSPEVVSANGPPTVKLYTSVENAQ 420
QY 421 QUKGTAIPHDINIGESPVVIGQYDNQHEQDRPIPSAPSPSPSVLKVANDVLMLSLTAAY 480
Db 421 QUKGTAIPHDINIGESPVVIGQYDNQHEQDRPIPSAPSPSPSVLKVANDVLMLSLTAAY 480
QY 481 DQSTYGSSTGPPVYSDSVTLVNVATGAQAVARSLDWTKVTLGDRPISTIQOYSKTFVLP 540
Db 481 DQSTYGSSTGPPVYSDSVTLVNVATGAQAVARSLDWTKVTLGDRPISTIQOYSKTFVLP 540
QY 541 LRGLSFEWAGCTTKAGYPYNYNTTASDQLLVENAAGRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRGLSFEWAGCTTKAGYPYNYNTTASDQLLVENAAGRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSAIAIEGHTIDYPARAHTFEDFPGCPPIGLQCCAFQSTVAELQBLKMKVKIREL 660
Db 601 LAPHSAIAIEGHTIDYPARAHTFEDFPGCPPIGLQCCAFQSTVAELQBLKMKVKIREL 660

Search completed: August 5, 2003, 09:27:03
Job time : 88 secs

GenCore version 5.1.6
Copyright (c) 1993-2003 CompuGen Ltd.

0M protein protein search, using sw model

Run on: August 5, 2003, 09:25:33 : Search time 30 seconds
(without alignments)
930.849 million cell updates/sec

111 res: US 09-851-410A-8
Post-test score: 34.06
Sequence: 1 MKPPEILLILLMLPMLPAP.....QSTVALQLKKVKGTREL 660

Scoring table: HESUM62

Gapop 10.0 , Gapext 9.5
Searched: 328717 seqs, 42410858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum hit seq length: 6

Maximum hit seq length: 200000000

Post processing: Minimum Match 100%

Listing first 45 summaries

Database: Issued_Patents_AA*
1: 5465_641 data/5465_641_5A_00MB.ppt*
2: 5465_641 data/5465_641_5B_00MB.ppt*
3: 5465_641 data/5465_641_5C_00MB.ppt*
4: 5465_641 data/5465_641_5D_00MB.ppt*
5: 5465_641 data/5465_641_5E_00MB.ppt*
6: 5465_641 data/5465_641_5F_00MB.ppt*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34.08	100.0	660	1	US-08-240-049B-15
2	34.08	100.0	660	1	US-08-259-148A-19
3	34.08	100.0	660	1	US-08-464-054-19
4	34.08	100.0	660	2	US-07-576-041A-19
5	34.08	100.0	660	3	US-08-478-507-8
6	34.08	100.0	660	3	US-08-542-644-13
7	34.08	100.0	660	3	US-09-128-275A-8
8	34.08	100.0	660	3	US-08-477-292-13
9	34.08	100.0	660	4	US-09-553-427-8
10	34.08	100.0	660	4	US-09-462-606-50
11	34.08	100.0	660	4	US-07-870-985A-19
12	34.08	100.0	660	5	PCT-US95-13703-13
13	33.99	99.5	660	4	US-09-172-699-2
14	33.91	99.5	660	4	US-09-462-606-12
15	33.82	99.2	660	4	US-08-840-416-2
16	33.82	99.2	660	3	US-08-859-523-2
17	33.82	99.2	660	3	US-08-471-471-2
18	33.82	99.2	660	4	US-09-462-606-49
19	33.82	99.2	660	4	US-09-462-606-56
20	33.82	99.2	660	4	US-09-402-775-2
21	33.82	99.2	660	5	PCT-US94-08649A-2
22	33.82	99.2	660	5	PCT-US93-08649-2
23	33.74	99.0	660	4	US-09-462-606-51
24	33.72	98.9	660	4	US-09-462-606-52
25	33.7	98.8	660	4	US-09-462-606-53
26	33.7	98.8	660	4	US-09-462-606-55
27	33.6	97.9	660	4	US-09-462-606-54

28	32.21	94.5	660	3	US-08-542-644-34	Sequence 14, Appl
29	32.21	94.5	660	3	US-08-542-644-34	Sequence 14, Appl
30	32.15	94.3	660	4	US-09-462-606-48	Sequence 49, Appl
31	31.94.5	93.7	659	1	US-08-240-049B-16	Sequence 15, Appl
32	31.94.5	93.7	659	1	US-08-259-148A-19	Sequence 19, Appl
33	31.94.5	93.7	659	1	US-08-464-054-20	Sequence 20, Appl
34	31.94.5	93.7	659	2	US-07-870-985A-20	Sequence 20, Appl
35	31.94.5	93.7	659	4	US-08-477-292-14	Sequence 14, Appl
36	31.94.5	93.7	659	4	US-07-870-985A-20	Sequence 20, Appl
37	31.64	92.8	660	4	US-09-462-606-57	Sequence 51, Appl
38	28.65	84.1	561	4	US-09-172-699-20	Sequence 20, Appl
39	28.21	82.8	552	4	US-09-172-699-16	Sequence 16, Appl
40	28.13	82.5	549	3	US-08-542-644-15	Sequence 15, Appl
41	28.13	82.5	549	3	US-08-477-292-15	Sequence 15, Appl
42	28.13	82.5	549	3	US-08-542-644-15	Sequence 15, Appl
43	27.69	81.2	540	3	US-08-542-644-25	Sequence 25, Appl
44	27.69	81.2	540	3	US-08-542-644-25	Sequence 25, Appl
45	26.94	79.0	525	3	US-08-542-644-27	Sequence 27, Appl

ADDITIONS

RESULT 1
US-08-240-049B-15
Sequence 15, Application US/08240049B
Patent No. 5,686,249
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R.
APPLICANT: Tam, Albert W.
APPLICANT: Yarbrough, Patricia O.
TITLE: A METHOD FOR DETECTING HIV-1 INFLUENZA VIRUS AND METHODS
NUMBER OF SEQUENCES: 21
COMPUTER PROGRAM: ALPHABET
ADDRESS: Genetech Technologies, Inc.
Street: 460 Point Street
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS DOS
SOFTWARE: Patent In Force #1,0, Version #1.0
CURRENT APPLICATION DATA:
APPLICANT NUMBER: 057-3/24-0, 049B
FILING DATE: 09 MAY 1994
CLASSIFICATION: 4.0
ATTORNEY/AGENT INFORMATION:
NAME: Charles K. Shott
REGISTRATION NUMBER: 08017
REFERENCE/KEYWORD NUMBER: 4,000,000, 4,000,000, 4,000,000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0960
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Report 15 + Virus (human strain)
INDIVIDUAL ISOLATE: ORF 2
US-08-240-049B-15

Query Match: 100.0% Score: 34.08 Length: 660
Best Local Similarity: 100.0% Prod. No.: 240-049B
Matches: 660 : Characteristic : 0 : Mismatches : 0 : Gaps : 0

QY 1 MKPPEILLILLMLPMLPAP.....QSTVALQLKKVKGTREL 660

1 FILING DATE: 17 JUN 1998
 2 APPLICANT: CHARM, YULIYU
 3 TITLE OF INVENTION: HEPATITIS B VIRUS ANTIGEN AND USES THEREOF
 4 NUMBER OF SEQUENCES: 41
 5 CORRESPONDENCE ADDRESS:
 6 ADDRESSEE: Tech Innov & Associates
 7 STREET: 450 Cambridge Avenue, Suite 200
 8 CITY: Palo Alto
 9 STATE: CA
 10 COUNTRY: USA
 11 ZIP: 94306
 12 COMPUTER READABLE FORM:
 13 MEDIUM TYPE: floppy disk
 14 COMPUTER: IBM pc compatible
 15 OPERATING SYSTEM: pc Dos/MS-DOS
 16 SOFTWARE: Patent In Police- # 1.0, Version # 1.2
 17 CURRENT APPLICATION DATA:
 18 APPLICATION NUMBER: US/09/429,634
 19 FILING DATE:
 20 CLASSIFICATION: 4A6
 21 ATTORNEY/AGENT INFORMATION:
 22 NAME: Ebbian, Gary B.
 23 REGISTRATION NUMBER: 43,875
 24 REFERENCE/LOCKET NUMBER: 400 0294, 40
 25 TELECOMMUNICATION INFORMATION:
 26 TELEPHONE: (415) 424 0880
 27 TELEFAX: (415) 424 0960
 28 INFORMATION FOR SEQ ID NO: 1:
 29 SEQUENCE CHARACTERISTICS:
 30 LENGTH: 660 amino acids
 31 TYPE: amino acid
 32 FEATURE: 1 linear
 33 MOLECULE TYPE: protein
 34 US 09 429 634 B

Query Match 100.0% Score: 408; ID #: Length: 660;
 Best Local Similarity 100.0% Ident. No. 2,40-294;
 Matches: 660; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 07 1 MPRPILLLLMFLMPLPAPPGGSGRGRGRSGSGRGGWGRVDSQPPAIPYIHTN 60
 10 1 MPRPILLLLMFLMPLPAPPGGSGRGRGRSGSGRGGWGRVDSQPPAIPYIHTN 60
 07 61 PPAPIVIAAGAGPRVOPAPPIAGSAGKQAGRAVASPRPRTTAGAAPTAVAPADTP 120
 10 61 PPAPIVIAAGAGPRVOPAPPIAGSAGKQAGRAVASPRPRTTAGAAPTAVAPADTP 120
 07 121 PPDVMSGCAILLRQYNISSDLSSTVATGNIIMLYAAPLSPLLPQGGTNTHTMATEAS 180
 10 121 PPDVMSGCAILLRQYNISSDLSSTVATGNIIMLYAAPLSPLLPQGGTNTHTMATEAS 180
 07 181 NYAGYVAVATIRYPLVFNAGGYAISISEWPTTITPTSDVMNSITSTIDRILVGGI 240
 10 181 NYAGYVAVATIRYPLVFNAGGYAISISEWPTTITPTSDVMNSITSTIDRILVGGI 240
 07 241 ASLIVPSRLHYRNQWHSVEISVAREEATISGLVMCIHSILVNSYNTIPYTGALL 400
 10 241 ASLIVPSRLHYRNQWHSVEISVAREEATISGLVMCIHSILVNSYNTIPYTGALL 400
 07 601 DPALEPRRLTGNRTGVSGYSTARRGRPRGAGTAAELTTTAAATKMKDLYETNG 560
 10 601 DPALEPRRLTGNRTGVSGYSTARRGRPRGAGTAAELTTTAAATKMKDLYETNG 560
 07 661 VREPRGIALIFENLAHTIAGIPTELLSSAGATFYSPVVSANRPTVKIYTSVNAQ 420
 10 661 VREPRGIALIFENLAHTIAGIPTELLSSAGATFYSPVVSANRPTVKIYTSVNAQ 420
 07 421 GIKGIAIPIHAGGIRVVGIGDGRGRGQAPFTISPAISKEFVLRANDVWLISLAAY 480
 10 421 GIKGIAIPIHAGGIRVVGIGDGRGRGQAPFTISPAISKEFVLRANDVWLISLAAY 480
 07 481 DASTYSSICPVVSVSVTLNVVAICAGAVAPSLIWTFTVTLGPPLSLQCYSKTFVLP 540
 10 481 DASTYSSICPVVSVSVTLNVVAICAGAVAPSLIWTFTVTLGPPLSLQCYSKTFVLP 540
 07 541 LPKRLTQWATIKAYIYVNTTASGGLVHNAAGHVAITVTTSLGAGVSTSAVAV 600
 10 541 LPKRLTQWATIKAYIYVNTTASGGLVHNAAGHVAITVTTSLGAGVSTSAVAV 600
 07 601 LAHSALALLELLIYPAPAHIFDGFCECHPLGLGCGAPSTVAIVGQPLMPKVPYTFEL 660
 10 601 LAHSALALLELLIYPAPAHIFDGFCECHPLGLGCGAPSTVAIVGQPLMPKVPYTFEL 660

RESULT 6
 US 09 429 634 14
 1 Sequence 14, Application US/09/429,634
 2 Patent No. 6,213,770
 3 GENERAL INFORMATION:
 4 APPLICANT: Ebbian, Gary B., Thomas K.
 5 ATTORNEY: McAtene, C. Patrick
 6 APPLICANT: Yarbough, Patricia G.

1 APPLICANT: Chann, YULIU
 2 TITLE OF INVENTION: HEPATITIS B VIRUS ANTIGEN AND USES THEREOF
 3 NUMBER OF SEQUENCES: 41
 4 CORRESPONDENCE ADDRESS:
 5 ADDRESSEE: Tech Innov & Associates
 6 STREET: 450 Cambridge Avenue, Suite 200
 7 CITY: Palo Alto
 8 STATE: CA
 9 COUNTRY: USA
 10 ZIP: 94306
 11 COMPUTER READABLE FORM:
 12 MEDIUM TYPE: floppy disk
 13 COMPUTER: IBM pc compatible
 14 OPERATING SYSTEM: pc Dos/MS-DOS
 15 SOFTWARE: Patent In Police- # 1.0, Version # 1.2
 16 CURRENT APPLICATION DATA:
 17 APPLICATION NUMBER: US/09/429,634
 18 FILING DATE:
 19 CLASSIFICATION: 4A6
 20 ATTORNEY/AGENT INFORMATION:
 21 NAME: Ebbian, Gary B.
 22 REGISTRATION NUMBER: 43,875
 23 REFERENCE/LOCKET NUMBER: 400 0294, 40
 24 TELECOMMUNICATION INFORMATION:
 25 TELEPHONE: (415) 424 0880
 26 TELEFAX: (415) 424 0960
 27 INFORMATION FOR SEQ ID NO: 1:
 28 SEQUENCE CHARACTERISTICS:
 29 LENGTH: 660 amino acids
 30 TYPE: amino acid
 31 FEATURE: 1 linear
 32 MOLECULE TYPE: protein
 33 HYPOTHETICAL: NO
 34 ORIGINAL SOURCE:
 35 INDIVIDUAL ISOLATE: Hepatitis B Virus (Burma strain)
 36 INDIVIDUAL ISOLATE: ORF 2
 37 US-08-542-634-14
 Query Match 100.0% Score: 408; ID #: Length: 660;
 Best Local Similarity 100.0% Ident. No. 2,40-294;
 Matches: 660; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 07 1 MPRPILLLLMFLMPLPAPPGGSGRGRGRSGSGRGGWGRVDSQPPAIPYIHTN 60
 10 1 MPRPILLLLMFLMPLPAPPGGSGRGRGRSGSGRGGWGRVDSQPPAIPYIHTN 60
 07 61 PPAPIVIAAGAGPRVOPAPPIAGSAGKQAGRAVASPRPRTTAGAAPTAVAPADTP 120
 10 61 PPAPIVIAAGAGPRVOPAPPIAGSAGKQAGRAVASPRPRTTAGAAPTAVAPADTP 120
 07 121 PPDVMSGCAILLRQYNISSDLSSTVATGNIIMLYAAPLSPLLPQGGTNTHTMATEAS 180
 10 121 PPDVMSGCAILLRQYNISSDLSSTVATGNIIMLYAAPLSPLLPQGGTNTHTMATEAS 180
 07 181 NYAGYVAVATIRYPLVFNAGGYAISISEWPTTITPTSDVMNSITSTIDRILVGGI 240
 10 181 NYAGYVAVATIRYPLVFNAGGYAISISEWPTTITPTSDVMNSITSTIDRILVGGI 240
 07 241 ASLIVPSRLHYRNQWHSVEISVAREEATISGLVMCIHSILVNSYNTIPYTGALL 400
 10 241 ASLIVPSRLHYRNQWHSVEISVAREEATISGLVMCIHSILVNSYNTIPYTGALL 400
 07 601 DPALEPRRLTGNRTGVSGYSTARRGRPRGAGTAAELTTTAAATKMKDLYETNG 560
 10 601 DPALEPRRLTGNRTGVSGYSTARRGRPRGAGTAAELTTTAAATKMKDLYETNG 560
 07 661 VREPRGIALIFENLAHTIAGIPTELLSSAGATFYSPVVSANRPTVKIYTSVNAQ 420
 10 661 VREPRGIALIFENLAHTIAGIPTELLSSAGATFYSPVVSANRPTVKIYTSVNAQ 420
 07 421 GIKGIAIPIHAGGIRVVGIGDGRGRGQAPFTISPAISKEFVLRANDVWLISLAAY 480
 10 421 GIKGIAIPIHAGGIRVVGIGDGRGRGQAPFTISPAISKEFVLRANDVWLISLAAY 480
 07 481 DASTYSSICPVVSVSVTLNVVAICAGAVAPSLIWTFTVTLGPPLSLQCYSKTFVLP 540
 10 481 DASTYSSICPVVSVSVTLNVVAICAGAVAPSLIWTFTVTLGPPLSLQCYSKTFVLP 540
 07 541 LPKRLTQWATIKAYIYVNTTASGGLVHNAAGHVAITVTTSLGAGVSTSAVAV 600
 10 541 LPKRLTQWATIKAYIYVNTTASGGLVHNAAGHVAITVTTSLGAGVSTSAVAV 600
 07 601 LAHSALALLELLIYPAPAHIFDGFCECHPLGLGCGAPSTVAIVGQPLMPKVPYTFEL 660
 10 601 LAHSALALLELLIYPAPAHIFDGFCECHPLGLGCGAPSTVAIVGQPLMPKVPYTFEL 660

```

QY 481 DQSYGSGTGPVYVSDVSVTLVNVATCAQAVAPSI DWTYKVTLDGRPISTIQQYSKTFFVLP 540
DB 481 DQSYGSSDGPVYVSDVSVTLVNVATCAQAVAPSI DWTYKVTLDGRPISTIQQYSKTFFVLP 540
QY 541 LRGKISFWFACTKAGYPYNYNTTASDQILVENAAGHRVAISTYTTSLDAGPVSLAVAV 600
DB 541 LRGKISFWFACTKAGYPYNYNTTASDQILVENAAGHRVAISTYTTSLDAGPVSLAVAV 600
QY 601 LAPHSALALLEDTLQYARAHIFDDFCPEKPIAGLQSCAFQSVIAELQRLKMKVKIRGL 660
DB 601 LAPHSALALLEDTLQYARAHIFDDFCPEKPIAGLQSCAFQSVIAELQRLKMKVKIRGL 660

```

RESULT 7

```

US-09-128-275A-8
: Sequence 8, Application US/09128275A
: Patent No. 6229005
: GENERAL INFORMATION:
: APPLICANT: Reyes, Gregory R
: APPLICANT: Yarbough, Patricia O
: APPLICANT: Bradley, Daniel W
: APPLICANT: Krawczynski, Krzysztof J
: APPLICANT: Tam, Albert
: APPLICANT: Fry, Kirk E
: TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dehlinger & Associates
: STREET: 350 Cambridge Avenue, Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US-09-128, 275A
: FILING DATE: 03-AUG-1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/279,823
: FILING DATE: 25-JUL-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/681,078
: FILING DATE: 05-APR-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/505,269
: FILING DATE: 05-APR-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/420,921
: FILING DATE: 13-OCT-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/367,486
: FILING DATE: 16-JUN-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/336,672
: FILING DATE: 11-APR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/208,997
: FILING DATE: 17-JUN-1988

```

```

: ATTORNEY/AGENT INFORMATION:
: NAME: Petithory, Joanne R.
: REGISTRATION NUMBER: 42,995
: REFERENCE/DOCCKET NUMBER: 4600-0183 24
: TELEPHONE: (650) 424-0980
: TELEFAX: (650) 324-0960
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:

```

```

: LENGTH: 660 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-128-275A 8

```

```

Query Match 100.0%; Score 3408; DB 3; Length 660;
Seq Local Similarity 100.0%; Pred. No. 2,48-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MRPPPLLLLLMLPLMPAPPQPSGRRKRRSGSGKGFWDQDRVDSQFFAIPYIHPTN 60
DB 1 MRPPPLLLLLMLPLMPAPPQPSGRRKRRSGSGKGFWDQDRVDSQFFAIPYIHPTN 60
QY 61 PFADVTAAAGACPPVPCPAPPICSAWPPNQAQPPAVASPPPPHTIAAADI TAVAPAHPTP 120
DB 61 PFADVTAAAGACPPVPCPAPPICSAWPPNQAQPPAVASPPPPHTIAAADI TAVAPAHPTP 120
QY 121 FVTEVSEKALIRQYMLSTPLTSSVATITRLVAVAAULSPLELQCTENIRKAMIEAS 180
DB 121 FVTEVSEKALIRQYMLSTPLTSSVATITRLVAVAAULSPLELQCTENIRKAMIEAS 180
QY 181 HVAAYAVABATRYPRVVDHAYVVAITGSMPTTETTYVYVYVYVYVYVYVYVYVY 240
DB 181 HVAAYAVABATRYPRVVDHAYVVAITGSMPTTETTYVYVYVYVYVYVYVYVYVY 240
QY 241 ASSTVYVPEPSTGPRQVADYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 300
DB 241 ASSTVYVPEPSTGPRQVADYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 300
QY 361 EFALPTETRMITPENTNTVSPASSTARRHPRGADRTAFLETTAATPRMKHLYFTSNG 360
DB 361 EFALPTETRMITPENTNTVSPASSTARRHPRGADRTAFLETTAATPRMKHLYFTSNG 360
QY 421 QDKGIALPHDIDGESKVVITQVNDHQEQRPITPSAPSKPTSVIPANIVIMLSIAAFY 480
DB 421 QDKGIALPHDIDGESKVVITQVNDHQEQRPITPSAPSKPTSVIPANIVIMLSIAAFY 480
QY 541 LRKLEFWAGCTKAYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 600
DB 541 LRKLEFWAGCTKAYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 600
QY 601 LAPHSALALLEDTLQYARAHIFDDFCPEKPIAGLQSCAFQSVIAELQRLKMKVKIRGL 660
DB 601 LAPHSALALLEDTLQYARAHIFDDFCPEKPIAGLQSCAFQSVIAELQRLKMKVKIRGL 660

```

RESULT 8

```

US-09-477-292-14
: Sequence 14, Application US/09477292
: Patent No. 6291641

```

```

: GENERAL INFORMATION:
: APPLICANT: Fuert, Thomas
: APPLICANT: McAttee, Patrick
: APPLICANT: Yarbough, Patricia
: APPLICANT: Zhang, Yifan
: TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Catherine M. Desjardins, M.D.
: STREET: 505 Penobscot Drive
: CITY: Redwood City
: STATE: CA
: COUNTRY: US
: ZIP: 94063

```



```

QY 121 EVVWVSGEALIEPQVNIESTPLTSVATGTHLVYAAVLSELLPLQSTRIHIMATKAS 180
DB 121 EVVWVSGEALIEPQVNIESTPLTSVATGTHLVYAAVLSELLPLQSTRIHIMATKAS 180
QY 141 NYAVSVARATIEYRIVFNAAVYAIPISEWPGTITPTISVMKSIETSHWPKIVQVQCI 240
DB 141 NYAVSVARATIEYRIVFNAAVYAIPISEWPGTITPTISVMKSIETSHWPKIVQVQCI 240
QY 241 ASPTVIFSEPIHYFNGQWISVTSQVAPTEASQVLMGTHPSI VNSVTNIPYTCALGIL 300
DB 241 ASPTVIFSEPIHYFNGQWISVTSQVAPTEASQVLMGTHPSI VNSVTNIPYTCALGIL 300
QY 301 DFALFIEPMLTIPNTNIPVSPYSSYSTARHPIEPGAGGTAFILTTAATSEIMKLYPTSTNG 360
DB 301 DFALFIEPMLTIPNTNIPVSPYSSYSTARHPIEPGAGGTAFILTTAATSEIMKLYPTSTNG 360
QY 361 VGTGEGTALTLENLAUTLGGTPTLTISSAGGQIFYSPPVVSANSEPTVKIYISVENAQ 420
DB 361 VGTGEGTALTLENLAUTLGGTPTLTISSAGGQIFYSPPVVSANSEPTVKIYISVENAQ 420
QY 421 QCKGATAPHITLGGESVWVQGVNQHGFGEPTTFSEAFSEPSVILKANIWLISIAAEY 480
DB 421 QCKGATAPHITLGGESVWVQGVNQHGFGEPTTFSEAFSEPSVILKANIWLISIAAEY 480
QY 481 DQSTYRSTGKVVSQSVILVNVAIAGAVAKSLIWKVILGKPLSLIAYSKIFVLP 540
DB 481 DQSTYRSTGKVVSQSVILVNVAIAGAVAKSLIWKVILGKPLSLIAYSKIFVLP 540
QY 541 LRKLSFWPACTKAGYPYNYNTASDQLLVENAAHRVAITSYTTSLCAGVTSISAVAV 600
DB 541 LRKLSFWPACTKAGYPYNYNTASDQLLVENAAHRVAITSYTTSLCAGVTSISAVAV 600
QY 601 LAPHSAIATLFTTIEYAPAHIEHDEPTEPQPIGLQGPAPSTIVAPLQPIKMKVQKTHPEI 660
DB 601 LAPHSAIATLFTTIEYAPAHIEHDEPTEPQPIGLQGPAPSTIVAPLQPIKMKVQKTHPEI 660

RESULT 13
US-09-172-699-2
: Sequence 2 Application us-09-172-699-2
: Patent No. 6514690
: GENERAL INFORMATION:
: APPLICANT: Anderson, David A.
: APPLICANT: Locarnini, Stephen A.
: APPLICANT: Torressi, Joseph
: APPLICANT: Hui, Zhuang
: APPLICANT: Li, Fan
: TITLE OF INVENTION: IMMUNOREACTIVE ANTIBODIES OF HEPATITIS E VIRUS
: FILE REFERENCE: Davies Col. Cave
: CURRENT APPLICATION NUMBER: US/099/172,699A
: PRIOR FILING DATE: 1998-10-14
: EARLIER APPLICATION NUMBERS: 02/013,727
: PRIOR FILING DATE: 1996-06-20
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 2
: LENGTH: 660
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-172-699-2

```

```

Query Match 99.7% Score 4399 DB 4 Length 660
Best Local Similarity 99.7% Prod. No. 1,50,293
Matches 658 Conservative 1 Mismatches 1 Indels 0 Gaps 0

```

```

QY 1 MRPEPILLILLMLPMLPAPHPQGPSRGRKSGSGSGSPWGLAVUSQFPAIFYTHTN 60
DB 1 MRPEPILLILLMLPMLPAPHPQGPSRGRKSGSGSGSPWGLAVUSQFPAIFYTHTN 60
QY 61 PPAHVTAAATATFETATLGEAMKSLGAPAVASERHPIIAVAAPLAVAPAHDP 120
DB 61 PPAHVTAAATATFETATLGEAMKSLGAPAVASERHPIIAVAAPLAVAPAHDP 120

```

```

QY 121 EVVWVSGEALIEPQVNIESTPLTSVATGTHLVYAAVLSELLPLQSTRIHIMATKAS 180
DB 121 EVVWVSGEALIEPQVNIESTPLTSVATGTHLVYAAVLSELLPLQSTRIHIMATKAS 180
QY 141 NYAVSVARATIEYRIVFNAAVYAIPISEWPGTITPTISVMKSIETSHWPKIVQVQCI 240
DB 141 NYAVSVARATIEYRIVFNAAVYAIPISEWPGTITPTISVMKSIETSHWPKIVQVQCI 240
QY 241 ASPTVIFSEPIHYFNGQWISVTSQVAPTEASQVLMGTHPSI VNSVTNIPYTCALGIL 300
DB 241 ASPTVIFSEPIHYFNGQWISVTSQVAPTEASQVLMGTHPSI VNSVTNIPYTCALGIL 300
QY 301 DFALFIEPMLTIPNTNIPVSPYSSYSTARHPIEPGAGGTAFILTTAATSEIMKLYPTSTNG 360
DB 301 DFALFIEPMLTIPNTNIPVSPYSSYSTARHPIEPGAGGTAFILTTAATSEIMKLYPTSTNG 360
QY 361 VGTGEGTALTLENLAUTLGGTPTLTISSAGGQIFYSPPVVSANSEPTVKIYISVENAQ 420
DB 361 VGTGEGTALTLENLAUTLGGTPTLTISSAGGQIFYSPPVVSANSEPTVKIYISVENAQ 420
QY 421 QCKGATAPHITLGGESVWVQGVNQHGFGEPTTFSEAFSEPSVILKANIWLISIAAEY 480
DB 421 QCKGATAPHITLGGESVWVQGVNQHGFGEPTTFSEAFSEPSVILKANIWLISIAAEY 480
QY 481 DQSTYRSTGKVVSQSVILVNVAIAGAVAKSLIWKVILGKPLSLIAYSKIFVLP 540
DB 481 DQSTYRSTGKVVSQSVILVNVAIAGAVAKSLIWKVILGKPLSLIAYSKIFVLP 540
QY 541 LRKLSFWPACTKAGYPYNYNTASDQLLVENAAHRVAITSYTTSLCAGVTSISAVAV 600
DB 541 LRKLSFWPACTKAGYPYNYNTASDQLLVENAAHRVAITSYTTSLCAGVTSISAVAV 600
QY 601 LAPHSAIATLFTTIEYAPAHIEHDEPTEPQPIGLQGPAPSTIVAPLQPIKMKVQKTHPEI 660
DB 601 LAPHSAIATLFTTIEYAPAHIEHDEPTEPQPIGLQGPAPSTIVAPLQPIKMKVQKTHPEI 660

RESULT 14
US-09-462-606-12
: Sequence 12 Application us-09-462-606-12
: Patent No. 6432408
: GENERAL INFORMATION:
: APPLICANT: MENG, XIANG-JIN
: APPLICANT: Emerson, Suzanne D.
: APPLICANT: Purcell, Robert H.
: TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
: FILE REFERENCE: 202402670S1
: CURRENT FILING DATE: 2000-06-12
: PRIOR APPLICATION NUMBER: US 60/053069
: PRIOR FILING DATE: 1997-07-18
: EARLIER APPLICATION NUMBERS: 02/014,665
: PRIOR FILING DATE: 1998-07-17
: NUMBER OF SEQ ID NOS: 65
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO: 12
: LENGTH: 660
: TYPE: PRT
: ORGANISM: Hepatitis E virus
US-09-462-606-12

```

```

Query Match 99.5% Score 4391 DB 4 Length 660
Best Local Similarity 99.4% Prod. No. 7,70,204
Matches 676 Conservative 1 Mismatches 3 Indels 0 Gaps 0

```

```

QY 1 MRPEPILLILLMLPMLPAPHPQGPSRGRKSGSGSGSPWGLAVUSQFPAIFYTHTN 60
DB 1 MRPEPILLILLMLPMLPAPHPQGPSRGRKSGSGSGSPWGLAVUSQFPAIFYTHTN 60
QY 61 PPAHVTAAATATFETATLGEAMKSLGAPAVASERHPIIAVAAPLAVAPAHDP 120
DB 61 PPAHVTAAATATFETATLGEAMKSLGAPAVASERHPIIAVAAPLAVAPAHDP 120

```

[illegible]

RESULTS 16,
US 08 840 316, 2
Application US/08840316

1 GENERAL INFORMATION:
2 APPLICANT: Emerson, Suzanne D., Porcely, Robert B.,
3 APPLICANT: Saeed, Saqiel A., and Robinson, Robin A.
4 TITLE OF INVENTION: Research on Folic Acid
5 TITLE OF INVENTION: A Pakistani Strain of Hepatitis E And Their
6 TITLE OF INVENTION: Use In Diagnostic Methods And Vaccines
7 NUMBER OF SEQUENCES: 11
8 CORRESPONDENT ADDRESS:
9 ADDRESSEE: MORGAN & FINNEGAN
10 STREET: 445 TAPPE AVENUE
11 CITY: NEW YORK
12 STATE: NEW YORK
13 COUNTRY: USA
14 ZIP: 10154
15 COMMUNICABLE FORM:
16 MEDIUM TYPE: FLOPPY DISK
17 COMPUTER: IBM PC COMPATIBLE
18 OPERATING SYSTEM: PC DOS/MS DOS
19 SOFTWARE: WORDPERFECT 5.1
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/040,316
22 FILING DATE: 11 APR 1997
23 CLASSIFICATION: 424
24 PRIORITY APPLICATION DATA:
25 APPLICATION NUMBER:
26 FILING DATE:
27 CLASSIFICATION: 424
28 AGENCY/AGENT INFORMATION:
29 NAME: Richard W. Bokk
30 REGISTRATION NUMBER: 46,459
31 PREFERRED/CAKEY NUMBER: 2726,4255
32 TELECOMMUNICATION INFORMATION:


```

QY 172 THIMATEASNYAGYVARATIRYRPLVNAVGVALSISEMPQTTTTPSVMSNISTD 241
DB 173 THIMATEASNYAGYVARATIRYRPLVNAVGVALSISEMPQTTTTPSVMSNISTD 120
QY 232 VRILVDPGIASELVIFSEPLHPNCGWPSVETSGVAEEFATSGVLMCHGSLVNSYNT 291
DB 233 VRILVDPGIASELVIFSEPLHPNCGWPSVETSGVAEEFATSGVLMCHGSLVNSYNT 180
QY 292 PYTGALGIDFALFIFEPNLTGNTNTPVSPSYSTAPHPPLPGGANGTAPLITTAATPMK 351
DB 293 PYTGALGIDFALFIFEPNLTGNTNTPVSPSYSTAPHPPLPGGANGTAPLITTAATPMK 240
QY 352 PLVFTSTNGVGFIPGIAITLLENIAITLGGFTLISASAGQGLEFYSFVWSAGCEPVK 411
DB 353 PLVFTSTNGVGFIPGIAITLLENIAITLGGFTLISASAGQGLEFYSFVWSAGCEPVK 400
QY 412 LYSVVENAQQKGAATPHILDGESRVVITQDYNGHQDPTTSTATSKFESVAFANIVL 471
DB 413 LYSVVENAQQKGAATPHILDGESRVVITQDYNGHQDPTTSTATSKFESVAFANIVL 290
QY 472 WLSTAAEYDQSYGSGTGPVVSQSVTLVNVATACAVASSLDKTKVTLGGRLSTLQQ 531
DB 473 WLSTAAEYDQSYGSGTGPVVSQSVTLVNVATACAVASSLDKTKVTLGGRLSTLQQ 420
QY 532 YSKTFTVLIRFKISEWFACTTKARYYNYNTTASDQIVENAAQSHVAISTYITSLAG 591
DB 533 YSKTFTVLIRFKISEWFACTTKARYYNYNTTASDQIVENAAQSHVAISTYITSLAG 480
QY 592 PVSISAVAVIAPHSAIAITLFTLDPAPAHIEFEPFEPFQDQSAQSLVAVHQAQPL 651
DB 593 PVSISAVAVIAPHSAIAITLFTLDPAPAHIEFEPFEPFQDQSAQSLVAVHQAQPL 540

```

RESULT 8

US-09-769-066-27

Sequence 27, Application US/09769066

Patent No. US20020107360A1

GENERAL INFORMATION:

APPLICANT: Fuerst, Thomas R.

McAttee, C. Patrick

Yarborough, Patricia O.

Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIBODIES AND USES THEREFOR

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Ave., Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09769066

FILING DATE: 24-Jan-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/542,634

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/SEQUENCE NUMBER: 4600-0293.30

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 525 amino acids

```

TYPE: amino acid
STRANDNESS: Hepatitis E virus (Barma strain)
762kDa, 56.5 kDa
Topology: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
SEQUENCE DESCRIPTION: SEQ ID NO: 27:
QC 09 769 066 27

```

```

Query Match 79.0% Score 2694. DP 10; Length 525;
Best Local Similarity 100.0%; Pred. No. 11e-216;
Matches 525; Concentrative 0; Mismatches 0; Gaps 0;

```

```

QY 112 AVFAHEHTFVTVISPTATLPPQYNLSPTSTSSVATGTNLVLYAAPLPLPLFVSTN 171
DB 113 AVFAHEHTFVTVISPTATLPPQYNLSPTSTSSVATGTNLVLYAAPLPLPLFVSTN 171
QY 146 1 AVFAHEHTFVTVISPTATLPPQYNLSPTSTSSVATGTNLVLYAAPLPLPLFVSTN 50
DB 147 1 AVFAHEHTFVTVISPTATLPPQYNLSPTSTSSVATGTNLVLYAAPLPLPLFVSTN 50
QY 171 THIMATEASNYAGYVARATIRYRPLVNAVGVALSISEMPQTTTTPSVMSNISTD 231
DB 172 THIMATEASNYAGYVARATIRYRPLVNAVGVALSISEMPQTTTTPSVMSNISTD 120
QY 232 VRILVDPGIASELVIFSEPLHPNCGWPSVETSGVAEEFATSGVLMCHGSLVNSYNT 291
DB 233 VRILVDPGIASELVIFSEPLHPNCGWPSVETSGVAEEFATSGVLMCHGSLVNSYNT 180
QY 292 PYTGALGIDFALFIFEPNLTGNTNTPVSPSYSTAPHPPLPGGANGTAPLITTAATPMK 351
DB 293 PYTGALGIDFALFIFEPNLTGNTNTPVSPSYSTAPHPPLPGGANGTAPLITTAATPMK 240
QY 352 PLVFTSTNGVGFIPGIAITLLENIAITLGGFTLISASAGQGLEFYSFVWSAGCEPVK 411
DB 353 PLVFTSTNGVGFIPGIAITLLENIAITLGGFTLISASAGQGLEFYSFVWSAGCEPVK 400
QY 412 LYSVVENAQQKGAATPHILDGESRVVITQDYNGHQDPTTSTATSKFESVAFANIVL 471
DB 413 LYSVVENAQQKGAATPHILDGESRVVITQDYNGHQDPTTSTATSKFESVAFANIVL 290
QY 472 WLSTAAEYDQSYGSGTGPVVSQSVTLVNVATACAVASSLDKTKVTLGGRLSTLQQ 531
DB 473 WLSTAAEYDQSYGSGTGPVVSQSVTLVNVATACAVASSLDKTKVTLGGRLSTLQQ 420
QY 532 YSKTFTVLIRFKISEWFACTTKARYYNYNTTASDQIVENAAQSHVAISTYITSLAG 591
DB 533 YSKTFTVLIRFKISEWFACTTKARYYNYNTTASDQIVENAAQSHVAISTYITSLAG 480
QY 592 PVSISAVAVIAPHSAIAITLFTLDPAPAHIEFEPFEPFQDQSAQSLVAVHQAQPL 651
DB 593 PVSISAVAVIAPHSAIAITLFTLDPAPAHIEFEPFEPFQDQSAQSLVAVHQAQPL 540

```

RESULT 9

US-09-769-066-16

Sequence 16, Application US/09769066

Patent No. US20020107360A1

GENERAL INFORMATION:

APPLICANT: Fuerst, Thomas R.

McAttee, C. Patrick

Yarborough, Patricia O.

Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIBODIES AND USES THEREFOR

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Ave., Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

1 SOFTWARE: Patent In Release #1.0, Version #1.25
 2 CURRENT APPLICATION DATA:
 3 APPLICATION NUMBER: US 09/066,066
 4 FILING DATE: 24 Jan 2001
 5 CLASSIFICATION: Unknown
 6 PRIOR APPLICATION DATA:
 7 APPLICATION NUMBER: Unknown
 8 FILING DATE: Unknown
 9 ATTORNEY/AGENT INFORMATION:
 10 NAME: Fabian, Gary R.
 11 REGISTRATION NUMBER: 43,875
 12 REFERENCE/AGENT INFORMATION:
 13 TELECOMMUNICATION INFORMATION:
 14 TELEPHONE: (415) 424-0680
 15 TELEFAX: (415) 424-0660
 16 INFORMATION FOR SEQ ID NO: 16:
 17 SEQUENCE CHARACTERISTICS:
 18 LENGTH: 549 amino acids
 19 TYPE: amino acid
 20 STRANDNESS: Linear
 21 MOLECULE TYPE: Protein
 22 HYPOTHETICAL: No
 23 ORIGINAL SOURCE:
 24 SEQUENCE DESCRIPTION: SEQ ID NO: 16:
 25

Query Match 72.99% Score 2684; ID 16; Length 549,
 Best Local Similarity 94.78; Prod. No. R.20-216;
 Matches 529; 5280 (94.1%) 14; Mismatches 15; Gaps 0.

27 112 AVAARDPPVDMISGAILRGYNLSSTPLSSVATGILNLYAAPISLLPLQDGTN 171
 28 1 AVAARDPPVDMISGAILRGYNLSSTPLSSVATGILNLYAAPISLLPLQDGTN 171
 29 120 IEMALANVAQVAVARATPYRGLVNAVGTATGTSWPTTTTTSVAMHETSTD 251
 30 1 IEMALANVAQVAVARATPYRGLVNAVGTATGTSWPTTTTTSVAMHETSTD 251
 31 1 IEMALANVAQVAVARATPYRGLVNAVGTATGTSWPTTTTTSVAMHETSTD 251
 32 242 VRLVQAGIASLVLTPSERLHYRNGWPSVETSAVAPFATSGVMCIHGSIVNSYNTI 291
 33 242 VRLVQAGIASLVLTPSERLHYRNGWPSVETSAVAPFATSGVMCIHGSIVNSYNTI 291
 34 241 VRLVQAGIASLVLTPSERLHYRNGWPSVETSAVAPFATSGVMCIHGSIVNSYNTI 291
 35 292 PYTGALGLDFALELEPNLTGNTNTRVSYSTFARHLPARGAIGFAELTTAAATPFMK 351
 36 1 PYTGALGLDFALELEPNLTGNTNTRVSYSTFARHLPARGAIGFAELTTAAATPFMK 351
 37 181 PYTGALGLDFALELEPNLTGNTNTRVSYSTFARHLPARGAIGFAELTTAAATPFMK 249
 38 542 DLYFTSNTAVDEFGALALFNLADHLAGLPTRELSSAGGQIFYSRPVVSANGPTVK 411
 39 241 DLYFTSNTAVDEFGALALFNLADHLAGLPTRELSSAGGQIFYSRPVVSANGPTVK 411
 40 412 LYSYVNAATGKTAIAFRTETLSESVVLELYRPHLEPTETSPATSPETVGLAHVL 471
 41 412 LYSYVNAATGKTAIAFRTETLSESVVLELYRPHLEPTETSPATSPETVGLAHVL 471
 42 412 LYSYVNAATGKTAIAFRTETLSESVVLELYRPHLEPTETSPATSPETVGLAHVL 471
 43 472 WSLGAAATGATGYSSTGTFVVSISVTLVHVAIAAQAARSLKWTAVTLGHELSLISG 541
 44 472 WSLGAAATGATGYSSTGTFVVSISVTLVHVAIAAQAARSLKWTAVTLGHELSLISG 541
 45 472 WSLGAAATGATGYSSTGTFVVSISVTLVHVAIAAQAARSLKWTAVTLGHELSLISG 541
 46 542 YSKTFVAGAGGKALFWALALFAGATFTRHLLASQGLVLRAAHRAVALTETLGAAG 591
 47 421 YSKTFVAGAGGKALFWALALFAGATFTRHLLASQGLVLRAAHRAVALTETLGAAG 591
 48 592 PYSIAVAVAPAPSAALATFETLQVAPARHPTGCPVPLGLGQAPGTSVAELPLER 651
 49 481 PYSIAVAVAPAPSAALATFETLQVAPARHPTGCPVPLGLGQAPGTSVAELPLER 651
 50 652 MKWGKTRRL 660
 51 541 VKWGRTRRL 549

RESULT 10
 US 09-769-066-26
 1 Sequence: 16, 411, 171, 291, 309/60066
 2 Patent No. US200207460A1
 3 GENERAL INFORMATION:
 4 APPLICANT: Postist, Thomas R.
 5 MEATON, C. Patrick
 6 Yaboum, Patrick O.
 7 Zhang, Yitun
 8 TITLE OF INVENTION: HERALDING VIRUS ANTIGENS AND USES THEREOF
 9 NUMBER OF SEQUENCES: 41
 10 CORRESPONDENCE ADDRESS:
 11 ADDRESSEE: Techpoint, S Associates
 12 STREET: 350 Cambridge Ave., Suite 250
 13 CITY: Palo Alto
 14 STATE: CA
 15 COUNTRY: USA
 16 ZIP: 94306
 17 COMPUTER READABLE FORM:
 18 MEDIUM TYPE: floppy disk
 19 COMPUTER: IBM PC compatible
 20 OPERATING SYSTEM: PC DOS/MS DOS
 21 SOFTWARE: Patent In Release #1.0, Version #1.25
 22 CURRENT APPLICATION DATA:
 23 APPLICATION NUMBER: 09/066,066
 24 FILING DATE: 24 Jan 2001
 25 CLASSIFICATION: Unknown
 26 PRIOR APPLICATION DATA:
 27 APPLICATION NUMBER: 08/542,634
 28 FILING DATE: Unknown
 29 ATTORNEY/AGENT INFORMATION:
 30 NAME: Fabian, Gary R.
 31 REGISTRATION NUMBER: 43,875
 32 REFERENCE/AGENT INFORMATION:
 33 TELECOMMUNICATION INFORMATION:
 34 TELEPHONE: (415) 424-0680
 35 TELEFAX: (415) 424-0660
 36 INFORMATION FOR SEQ ID NO: 26:
 37 SEQUENCE CHARACTERISTICS:
 38 LENGTH: 540 amino acids
 39 TYPE: amino acid
 40 STRANDNESS: Hepatitis E virus (Mexico strain)
 41 102KDa, 6.1 kDa
 42 Topology: linear
 43 MOLECULE TYPE: protein
 44 HYPOTHETICAL: No
 45 ORIGINAL SOURCE:
 46 SEQUENCE DESCRIPTION: SEQ ID NO: 26:
 47

Query Match 77.69% Score 2644; ID 16; Length 540;
 Best Local Similarity 94.88; Prod. No. 1.00-212;
 Matches 512; Gaps 14; Mismatches 15; Indels 0; Gaps 0;
 27 112 AVAARDPPVDMISGAILRGYNLSSTPLSSVATGILNLYAAPISLLPLQDGTN 171
 28 1 AVAARDPPVDMISGAILRGYNLSSTPLSSVATGILNLYAAPISLLPLQDGTN 171
 29 172 IEMALANVAQVAVARATPYRGLVNAVGTATGTSWPTTTTTSVAMHETSTD 251
 30 1 IEMALANVAQVAVARATPYRGLVNAVGTATGTSWPTTTTTSVAMHETSTD 251
 31 1 IEMALANVAQVAVARATPYRGLVNAVGTATGTSWPTTTTTSVAMHETSTD 251
 32 242 VRLVQAGIASLVLTPSERLHYRNGWPSVETSAVAPFATSGVMCIHGSIVNSYNTI 291
 33 242 VRLVQAGIASLVLTPSERLHYRNGWPSVETSAVAPFATSGVMCIHGSIVNSYNTI 291
 34 241 VRLVQAGIASLVLTPSERLHYRNGWPSVETSAVAPFATSGVMCIHGSIVNSYNTI 291
 35 292 PYTGALGLDFALELEPNLTGNTNTRVSYSTFARHLPARGAIGFAELTTAAATPFMK 351
 36 1 PYTGALGLDFALELEPNLTGNTNTRVSYSTFARHLPARGAIGFAELTTAAATPFMK 351
 37 181 PYTGALGLDFALELEPNLTGNTNTRVSYSTFARHLPARGAIGFAELTTAAATPFMK 249
 38 542 DLYFTSNTAVDEFGALALFNLADHLAGLPTRELSSAGGQIFYSRPVVSANGPTVK 411
 39 241 DLYFTSNTAVDEFGALALFNLADHLAGLPTRELSSAGGQIFYSRPVVSANGPTVK 411
 40 412 LYSYVNAATGKTAIAFRTETLSESVVLELYRPHLEPTETSPATSPETVGLAHVL 471
 41 412 LYSYVNAATGKTAIAFRTETLSESVVLELYRPHLEPTETSPATSPETVGLAHVL 471
 42 412 LYSYVNAATGKTAIAFRTETLSESVVLELYRPHLEPTETSPATSPETVGLAHVL 471
 43 472 WSLGAAATGATGYSSTGTFVVSISVTLVHVAIAAQAARSLKWTAVTLGHELSLISG 541
 44 472 WSLGAAATGATGYSSTGTFVVSISVTLVHVAIAAQAARSLKWTAVTLGHELSLISG 541
 45 472 WSLGAAATGATGYSSTGTFVVSISVTLVHVAIAAQAARSLKWTAVTLGHELSLISG 541
 46 542 YSKTFVAGAGGKALFWALALFAGATFTRHLLASQGLVLRAAHRAVALTETLGAAG 591
 47 421 YSKTFVAGAGGKALFWALALFAGATFTRHLLASQGLVLRAAHRAVALTETLGAAG 591
 48 592 PYSIAVAVAPAPSAALATFETLQVAPARHPTGCPVPLGLGQAPGTSVAELPLER 651
 49 481 PYSIAVAVAPAPSAALATFETLQVAPARHPTGCPVPLGLGQAPGTSVAELPLER 651
 50 652 MKWGKTRRL 660
 51 541 VKWGRTRRL 549

```

UY 412 YTSVFNACQFFCAIHEICLSEFFVWVYTYNDEHFESEPTSPAPSEFFFAVLEAVY 477
DB 301 YTSVFNACQKGAIDPHDINDSPRVVIGQVYHQBQPPPTSPAPSEFFFAVLEAVY 460
UY 472 WLSLTAAYEYDQSTYSSSTFVYVSUSVTIWSVAFCAAVAKPSIDWPKVTLQKPLSTLQK 531
DB 361 WLSLTAAYEYDQSTYSSSTFVYVSUSVTIWSVAFCAAVAKPSIDWPKVTLQKPLSTLQK 470
UY 532 YSKTFVPIPIRGKLSKFWPACTTKAGYPYNYNTTASQGLVENAACHPVAISTYTTSLQAG 591
DB 421 YSKTFVPIPIRGKLSKFWPACTTKAGYPYNYNTTASQGLVENAACHPVAISTYTTSLQAG 480
UY 592 PVSISAVAVLAPHSALALLEDTEDYDGRAHIFUDGCEKALGIDQCAPUSVAELQPLK 540
DB 481 PVSISAVAVLAPHSALALLEDTEDYDGRAHIFUDGCEKALGIDQCAPUSVAELQPLK 540

```

RESULT 11

US-09-769-066-38

Sequence 28, Application US/09769066

Patent No. US20020107360A1

GENERAL INFORMATION:

APPLICANT: Fuertes, Thomas R.

McAtee, C. Patrick

Yarbough, Patricia O.

Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Ave., Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/769,066

FILING DATE: 24 Jan 2001

CLASSIFICATION: <Unknown>

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/542,634

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 4600-0293.30

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 525 amino acids

TYPE: amino acid

STRANDEDNESS: Hepatitis E virus (Mexico strain)

r62kDa, 56.5 kDa

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

SEQUENCE DESCRIPTION: SEQ ID NO: 28:

US-09-769-066-38

```

Query Match: 75.4% Score 2559; DB 10; Length 525;
Best Local Similarity 94.7%; Proc. No. 3.2c 205;
Matches 497; Conservative 13; Mismatches 15; Indels 0; Caps 0.
UY 112 AVAPAHPTPPVDPVDSGAILRPGYNLSPLTSSVATGNTLVLYAAPLSLPLQDQIN 171

```

```

DB 1 AVAPAHPTPPVDPVDSGAILRPGYNLSPLTSSVATGNTLVLYAAPLSLPLQDQIN 400
UY 172 THIMATLAAAYAYVAFATFEYFPIVPAVAGVAVATGSMWNSFTTPTFVQWPPVQDPP 421
DB 61 THIMATLAAAYAYVAFATFEYFPIVPAVAGVAVATGSMWNSFTTPTFVQWPPVQDPP 420
UY 237 VFILVQRIASHIVITSEPIHYSEPIWESVFIQSWAFIATGIVMCTGNSVAVQVAVT 291
DB 121 VFILVQRIASHIVITSEPIHYSEPIWESVFIQSWAFIATGIVMCTGNSVAVQVAVT 180
UY 292 FYTALCLLLELLELLELLELLELLELLELLELLELLELLELLELLELLELLELLEL 451
DB 191 FYTALCLLLELLELLELLELLELLELLELLELLELLELLELLELLELLELLEL 240
UY 352 LAYFISINWSEFGPGLALIFNADIIQSDGDFECSAGSDGYSQWVAVANGPQVAVK 411
DB 241 LAYFISINWSEFGPGLALIFNADIIQSDGDFECSAGSDGYSQWVAVANGPQVAVK 400
UY 412 YTSVFNACQFFCAIHEICLSEFFVWVYTYNDEHFESEPTSPAPSEFFFAVLEAVY 471
DB 301 YTSVFNACQKGAIDPHDINDSPRVVIGQVYHQBQPPPTSPAPSEFFFAVLEAVY 460
UY 472 WLSLTAAYEYDQSTYSSSTFVYVSUSVTIWSVAFCAAVAKPSIDWPKVTLQKPLSTLQK 531
DB 361 WLSLTAAYEYDQSTYSSSTFVYVSUSVTIWSVAFCAAVAKPSIDWPKVTLQKPLSTLQK 470
UY 532 YSKTFVPIPIRGKLSKFWPACTTKAGYPYNYNTTASQGLVENAACHPVAISTYTTSLQAG 591
DB 421 YSKTFVPIPIRGKLSKFWPACTTKAGYPYNYNTTASQGLVENAACHPVAISTYTTSLQAG 480
UY 592 PVSISAVAVLAPHSALALLEDTEDYDGRAHIFUDGCEKALGIDQCAPUSVAELQPLK 540
DB 481 PVSISAVAVLAPHSALALLEDTEDYDGRAHIFUDGCEKALGIDQCAPUSVAELQPLK 540

```

RESULT 12

US-09-769-066-17

Sequence 17, Application US/09769066

Patent No. US20020107360A1

GENERAL INFORMATION:

APPLICANT: Fuertes, Thomas R.

McAtee, C. Patrick

Yarbough, Patricia O.

Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Ave., Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 08/542,634

FILING DATE: 24 Jan 2001

CLASSIFICATION: <Unknown>

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/542,634

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 4600-0293.30

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

? INFORMATION FOR SEQ ID NO: 17:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 427 amino acids
 ? STRAIN/DISEASE: Hepatitis E Virus (Horma strain) SG3
 ? FUNCTION: Replication
 ? MOLECULAR TYPE: Linear
 ? HYDROPHILIC No.
 ? ORIGINAL SOURCE:
 ? SEQUENCE DESCRIPTION: SEQ ID NO: 17:
 US 09 769 066-17

Query Match 49.2% Score 1677; DB 10; Length 327;
 Best Local Similarity 100.0%; Pred. No. 5,96-132;
 Matches 427; Conservative 0; Mismatches 0; Labels 0; Gaps 0;
 QY 43 LRRGALGTAELTIAATREMKHATISINWGEGRGIALLPNALDLSGLPFIETSSAGG 493
 DB 1 LAGGTAAELTIAATREMKHATISINWGEGRGIALLPNALDLSGLPFIETSSAGG 69
 QY 44 LPPYSRPVVSAN-PTTVLYISVNA-EE-K-TAIPR-DEG-GE-SHVVVIGQYINQHEPFI 453
 DB 61 LPPYSRPVVSAN-PTTVLYISVNA-EE-K-TAIPR-DEG-GE-SHVVVIGQYINQHEPFI 129
 QY 45 PSPAPSPSPSVLRANIVLWLSITAAFYQSTVGSSTGPPVYVSQSVTLVNVATGACAVAPS 513
 DB 21 PSPAPSPSPSVLRANIVLWLSITAAFYQSTVGSSTGPPVYVSQSVTLVNVATGACAVAPS 180
 QY 46 LHWKTVLGGRLSTLQYSPFIIVDP-ET-SWEA-ETPA-VYVYNNNTASTGGLIVEN 573
 DB 21 LHWKTVLGGRLSTLQYSPFIIVDP-ET-SWEA-ETPA-VYVYNNNTASTGGLIVEN 240
 QY 47 A-GRVAISITYITSLGAGPVSISAVAVLAPHSALALPDLTLYPARAHITFDGFCGCRPL 633
 DB 241 A-GRVAISITYITSLGAGPVSISAVAVLAPHSALALPDLTLYPARAHITFDGFCGCRPL 400
 QY 48 GLEQAFQSTVAELGRLKMKVGRKPEL 660
 DB 501 GLEQAFQSTVAELGRLKMKVGRKPEL 427

RESULT 14
 US-09 468 147-206
 ? Sequence 207; Application US/09468147A
 ? Publication No. US20040049601A1
 ? GENERAL INFORMATION:
 ? APPLICANT: Abbott Laboratories
 ? APPLICANT: Schlader, George G.
 ? APPLICANT: Eker, James C.
 ? APPLICANT: Desai, Suresh M.
 ? APPLICANT: Dawson, George J.
 ? APPLICANT: Mushahwar, T. K.
 ? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
 ? FILE REFERENCE: 6242-US-P1
 ? CURRENT APPLICATION NUMBER: US/09/468,147A
 ? FILING DATE: 1999-12-21
 ? EARLIER APPLICATION NUMBER: US 09/174,141
 ? EARLIER FILING DATE: 1998-10-15
 ? EARLIER APPLICATION NUMBER: US 60/061,199
 ? EARLIER FILING DATE: 1997-10-15
 ? NUMBER OF SEQ ID NOS: 258
 ? SOFTWARE: FASTSEQ for Windows Version 4.0
 ? SEQ ID NO: 206
 ? LENGTH: 459
 ? TYPE: PFI
 ? ORGANISM: Hepatitis E Virus
 ? FEATURES:
 ? OTHER INFORMATION: CKS-RF42M 4.p.p.
 US 09 468 147-206

Query Match 47.4% Score 1615; DB 11; Length 459;
 Best Local Similarity 95.6%; Pred. No. 1,56-126;
 Matches 408; Conservative 14; Mismatches 7; Labels 0; Gaps 0;

Best Local Similarity 93.6%; Pred. No. 1,56-126;
 Matches 408; Conservative 14; Mismatches 7; Labels 0; Gaps 0;
 QY 43 LRRGALGTAELTIAATREMKHATISINWGEGRGIALLPNALDLSGLPFIETSS 490
 DB 122 LRRGALGTAELTIAATREMKHATISINWGEGRGIALLPNALDLSGLPFIETSS 181
 QY 44 LPPYSRPVVSAN-PTTVLYISVNA-EE-K-TAIPR-DEG-GE-SHVVVIGQYINQHEPFI 450
 DB 162 LPPYSRPVVSAN-PTTVLYISVNA-EE-K-TAIPR-DEG-GE-SHVVVIGQYINQHEPFI 241
 QY 45 PSPAPSPSPSVLRANIVLWLSITAAFYQSTVGSSTGPPVYVSQSVTLVNVATGACAV 510
 DB 242 PSPAPSPSPSVLRANIVLWLSITAAFYQSTVGSSTGPPVYVSQSVTLVNVATGACAV 401
 QY 46 LHWKTVLGGRLSTLQYSPFIIVDP-ET-SWEA-ETPA-VYVYNNNTASTGGL 570
 DB 492 LHWKTVLGGRLSTLQYSPFIIVDP-ET-SWEA-ETPA-VYVYNNNTASTGGL 401
 QY 47 A-GRVAISITYITSLGAGPVSISAVAVLAPHSALALPDLTLYPARAHITFDGFC 630
 DB 92 LKSGA-GRVAISITYITSLGAGPVSISAVAVLAPHSALALPDLTLYPARAHITFDGFC 411
 QY 48 GLEQAFQSTVAELGRLKMKVGRKPEL 660
 DB 422 GLEQAFQSTVAELGRLKMKVGRKPEL 450

RESULT 14
 US-09 468 147-207
 ? Sequence 207; Application US/09468147A
 ? Publication No. US20040049601A1
 ? GENERAL INFORMATION:
 ? APPLICANT: Abbott Laboratories
 ? APPLICANT: Schlader, George G.
 ? APPLICANT: Eker, James C.
 ? APPLICANT: Desai, Suresh M.
 ? APPLICANT: Dawson, George J.
 ? APPLICANT: Mushahwar, T. K.
 ? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
 ? FILE REFERENCE: 6242-US-P1
 ? CURRENT APPLICATION NUMBER: US/09/468,147A
 ? FILING DATE: 1999-12-21
 ? EARLIER APPLICATION NUMBER: US 09/174,141
 ? EARLIER FILING DATE: 1998-10-15
 ? EARLIER APPLICATION NUMBER: US 60/061,199
 ? EARLIER FILING DATE: 1997-10-15
 ? NUMBER OF SEQ ID NOS: 258
 ? SOFTWARE: FASTSEQ for Windows Version 4.0
 ? SEQ ID NO: 207
 ? LENGTH: 459
 ? TYPE: PFI
 ? ORGANISM: Hepatitis E Virus
 ? FEATURES:
 ? OTHER INFORMATION: PLO-RF42M 14 3.p.p.
 US-09 468 147-207

Query Match 47.4% Score 1615; DB 11; Length 459;
 Best Local Similarity 95.6%; Pred. No. 1,56-126;
 Matches 408; Conservative 14; Mismatches 7; Labels 0; Gaps 0;
 QY 43 LRRGALGTAELTIAATREMKHATISINWGEGRGIALLPNALDLSGLPFIETSS 490
 DB 122 LRRGALGTAELTIAATREMKHATISINWGEGRGIALLPNALDLSGLPFIETSS 181
 QY 44 LPPYSRPVVSAN-PTTVLYISVNA-EE-K-TAIPR-DEG-GE-SHVVVIGQYINQHEPFI 450
 DB 162 LPPYSRPVVSAN-PTTVLYISVNA-EE-K-TAIPR-DEG-GE-SHVVVIGQYINQHEPFI 241
 QY 45 PSPAPSPSPSVLRANIVLWLSITAAFYQSTVGSSTGPPVYVSQSVTLVNVATGACAV 510
 DB 242 PSPAPSPSPSVLRANIVLWLSITAAFYQSTVGSSTGPPVYVSQSVTLVNVATGACAV 401

Search completed: August 5, 2003, 09:31:42
Jed film: 27 secs

```
QY 511 APSINWTVLDDPRLSTIGCYSKIEFVLHPZELSFWEAGTTEACYPYNYNTASDQIL 570
DB 302 ARSLTWSKVLTDGRLPTTIOQYKSTFYVLPURKLSFWEAGTTKAGYPYNYNTASDQIL 361
QY 571 VENAAGHRVAISTVTSIGAGPVSISAVAVLAPHSALALEDTLDYPARAHFTDDPCPC 630
DB 462 TENAAGHRVAISTVTSIGAGPVSISAVGVLAIPHSAALAVLEDTLDYPARAHFTDDPCPC 421
QY 641 RPLGLQCAFQSTVAELQRLKMKVGKTR 659
DB 422 RTLAGLQCAFQSTIAELQRLKMKVGKTR 450
```

RESULT 15

```
US-09-468-147-199
: Sequence 199, Application US/09468147A
: Publication No. US20030049601A1
: GENERAL INFORMATION:
: APPLICANT: Abbott Laboratories
: APPLICANT: Schlauder, George G.
: APPLICANT: Erker, James C.
: APPLICANT: Desai, Suresh M.
: APPLICANT: Dawson, George J.
: APPLICANT: Mushahwar, I. K.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
: TITLE OF INVENTION: HEPATITIS E VIRUS
: FILE REFERENCE: 6242 US.P1
: CURRENT APPLICATION NUMBER: US/09/468,147A
: CURRENT FILING DATE: 1999-12-21
: EARLIER APPLICATION NUMBER: US 09/173,141
: EARLIER FILING DATE: 1998-10-15
: EARLIER APPLICATION NUMBER: US 60/061,199
: EARLIER FILING DATE: 1997-10-15
: NUMBER OF SEQ ID NOS: 258
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 199
: LENGTH: 338
: TYPE: PRT
: ORGANISM: Hepatitis E Virus
: FEATURE:
: OTHER INFORMATION: cksort2m-2.pep
US-09-468-147-199
```

Query Match 47.0%; Score 1601; DB 11; Length 338;
Best Local Similarity 93.6%; Pred. No. 1.4e-125;
Matches 305; Conservative 14; Mismatches 7; Indels 0; Gaps 0;

```
QY 334 GADGTAEITTTAATRPKKDLYFTSTNGVGHIGRGIALTFNLAVLLGGLPTELISSAGG 393
DB 4 GADGTAEITTTAATRPKKDLHFAGTNGVGHVGRGIALTFNLADTVLLGGLPTELISSAGG 63
QY 454 QLEYSRPVVSANGPTVKLYTSVENAGQDKGIALPHIDLGESFVVLQDYLNHHEQDPT 453
DB 64 QLEYSRPVVSANGPTVKLYTSVENAGQDKGIIIPHIDLGESFVVLQDYLNHHEQDPT 123
QY 454 PSPAPSPFVSRLANDVWLSLTAAEYDQSTYGSSTGCVVVSVDVTLVAVATCAQAVARS 513
DB 124 PSPAPSPFVSRLANDVWLSLTAAEYDQSTYGSSTGCVVVSVDVTLVAVATCAQAVARS 183
QY 514 LDWTKVTLDPPISTIOQYKSTFYVLPURKLSFWEAGTTKAGYPYNYNTASDQIL 573
DB 184 LDWTKVTLDPPISTIOQYKSTFYVLPURKLSFWEAGTTKAGYPYNYNTASDQIL 243
QY 574 AAGHRVAISTVTSIGAGPVSISAVAVLAPHSALALEDTLDYPARAHFTDDPCPC 633
DB 244 AAGHRVAISTVTSIGAGPVSISAVGVLAIPHSAALAVLEDTLDYPARAHFTDDPCPC 303
QY 634 GLQSCAFQSTVAELQRLKMKVGKTR 659
DB 304 GLQSCAFQSTIAELQRLKMKVGKTR 329
```

Genome version 5.1.6
Copyright (c) 1994-2003 Computer Ltd.

OM protein protein search, using sw model

Run on: August 5, 2003, 09:24:33 ; Search time 44 Seconds
(without alignments)
1442.530 Million cell updates/sec

Title: US 09 851 410A-8
Post hoc score: 3.08
Sequence: 1 MKRPRLILLIMLPLPAP.....QSTVARLGRLEKPVGTREL 660

Search table: h080962
Gapop 10.0 ; Gapext 0.5

Searched: 28308 seqs, 9676682 residues

Total number of hits satisfying chosen parameters: 28308

Minimum hit seq length: 0

Maximum hit seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

hit base : PK_26: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44.98	100.0	669	1	VHWH2	structural protein
2	41.94	98.7	659	1	B44212	structural protein
3	14.3	4.2	1041	2	S55862	probable membrane
4	13.5	4.1	5188	2	B85547	probable RIX family
5	13.5	4.1	5291	2	F90646	hypothetical prote
6	13.7	4.0	3570	2	F35025	macin MOC5B, trach
7	13.5	4.0	1424	2	S52863	UNA-binding protei
8	13.5	4.0	4282	2	E82750	hemagglutinin-like
9	13.5	4.0	4442	2	E82589	hemagglutinin-like
10	13.5	4.0	4455	2	B82519	hemagglutinin-like
11	13.4	3.9	795	2	L21450	hypothetical prote
12	13.3	3.9	1802	2	S67903	HRF protein prote
13	13.5	3.9	583	2	S67571	hypothetical prote
14	13.5	3.8	1104	2	S59310	probable membrane
15	13.5	3.8	1161	2	S57180	probable membrane
16	13.5	3.8	1537	2	S53455	thrombulation prote
17	13.5	3.8	4013	2	A80480	probable invasin Y
18	12.9	3.8	635	2	F75477	hypothetical prote
19	12.9	3.8	1367	1	S48478	glucan 1,4-alpha-q
20	12.9	3.8	2130	2	A80825	probable exported
21	12.9	3.8	918	2	T02759	hypothetical prote
22	12.9	3.8	1515	2	S55062	phagocin precursor
23	12.9	3.8	1467	2	S51959	hypothetical prote
24	12.6	3.8	767	1	J06474	glucan 1,4-alpha-q
25	12.8	3.8	778	1	ALBYG	glucan 1,4-alpha-q
26	12.8	3.8	832	2	B87674	AHC transporter, H
27	12.7	3.7	518	1	F20147	L2 protein - human
28	12.7	3.7	400	2	T02852	probable membrane
29	12.7	3.7	460	2	T14650	tail fiber protein

30	127.5	3.7	978	2	114308	phage lambda relat
31	127	3.7	1275	2	133009	hypothetical prote
32	127	3.7	7463	2	130238	clA peptide synthet
33	126.5	3.7	351	1	VVELE	coat protein - tat
34	126.5	3.7	948	2	111678	hypothetical prote
35	126.5	3.7	1306	2	S25306	MSB2 protein - yoa
36	126	3.7	1037	2	113050	transcription fact
37	125.5	3.7	725	2	A41258	extracellular enzy
38	125.5	3.7	1879	2	S74915	hypothetical prote
39	125.5	3.7	1952	2	148814	hypothetical prote
40	124	3.6	963	2	140240	hypothetical prote
41	124	3.6	1131	2	141144	hypothetical serin
42	124	3.6	1483	2	C97052	probably cell fusio
43	123.5	3.6	797	1	V94931	glycyl-stere X pro
44	123	3.6	2242	2	144434	hypothetical prote
45	122.5	3.6	1324	2	134650	peptide synthetase

ALIGNMENTS

RESULT 1

VHWH2
structural protein 2 precursor hepatitis E virus (strain Burma)
C:Species: hepatitis E virus
C:Date: 30-Sep-1992 #sequence revision 30-Sep-1992 #text change 2-Jul-1999
C:Accession: C40778
R:Ram, A.W.; Smith, M.M.; Guerra, M.E.; Bann, C.C.; Bradley, D.W.; Fry, K.E.; Reyes, Virology 185, 120-141, 1991
A:Title: Hepatitis E virus (HEV): molecular cloning and sequencing of the full length
A:Reference number: A40778; M010:92024067; IM10:192670
A:Accession: C40778
A:Molecule type: genomic RNA
A:Residues: 1 660 -TAM-
A:Cross-references: G5074218; NID:430025; FID:AAA45061; FID:430029
A:Note: the authors translated the codon cap for residue 2 as Ala
C:Superfamily: hepatitis E virus structural protein 2
C:Keywords: structural protein
F:1 22/Domain: signal sequence #status predicted -SP1
F:23 660/Product: structural protein 2 #status predicted -SP2

Query Match	100.0%	Score	34.082	DB 1	Length	660	
Best Local Similarity	100.0%	Prod. No.	660				
Matches	660	Conservative	0	Mismatches	0	Gaps	0
Q7	1	MEPRRLILLIMLPLPAP.....QSTVARLGRLEKPVGTREL 660					
1b	1	MEPRRLILLIMLPLPAP.....QSTVARLGRLEKPVGTREL 660					
Q7	61	PLATVTAAGAGRVKQAFKLSAMHQAQRTAVASREFFETASAAATAVATAP 120					
1b	61	PLATVTAAGAGRVKQAFKLSAMHQAQRTAVASREFFETASAAATAVATAP 120					
Q7	121	LVTVVSEKALIKQYRLKELKELKELKELKELKELKELKELKELKELKEL 180					
1b	121	LVTVVSEKALIKQYRLKELKELKELKELKELKELKELKELKELKELKEL 180					
Q7	181	RYACYAVARATRYKRVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAV 240					
1b	181	RYACYAVARATRYKRVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAV 240					
Q7	241	ASEVATSEKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRY 300					
1b	241	ASEVATSEKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRY 300					
Q7	301	PLAELPERNLPPNTRKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRY 360					
1b	301	PLAELPERNLPPNTRKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRY 360					
Q7	361	VOELGKGLIIFNLARILLGLPTELISSAVVGLFYSEIVVAVNPEPVKLY 420					
1b	361	VOELGKGLIIFNLARILLGLPTELISSAVVGLFYSEIVVAVNPEPVKLY 420					

DB	2866	PARTAIITVTI-NGONTSATIDAGNWS-VTVPASVASALGEATYSV-ASVTNAGNSSTA	2923
QY	548	---WEAGTTKAGPYPNYNTTASQLLVENAAGHPVAISTYTTISLGA	591
DB	2924	SHNYGVN-IALP3--ITINPVATGDIINASEAGSAOTLSCQVIGAAAG	2968

RESILLIT 6

T45025
 mucin MUC5B, tracheobronchial [Imported] - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T45025
 R:Dessevry, J.L., Guyonnet-Dupont, V., Forchet, R., Aubert, J.F., Laine, A.
 J. Biol. Chem. 272, 3168-3178, 1997
 A:Title: Human mucin gene MUC5B, the 10.7 kb large central exon encodes various alternative splicing products
 A:Reference number: 272869; MIM:607451; PMID:931350
 A:Accession: T45025
 A:Status: Preliminary; translated from CB/EMBL/GenBank
 A:Molecule type: DNA
 A:Features: 1370 CDS
 A:Cross-references: EMBL:Z72436; NID:gl834502; PID:CAA96577.1; PID:gl834503
 A:Experimental source: Placenta
 C:Genetics:
 A:Gene: MUC5B

Query Match 4 0% Score 127, DP 2, Length 470.
Best Local Similarity 20.1%, Field No. 6.7,
Matches 143, Conservative 92, Mismatches 257, Indels 240, Gaps 32.

QY	52	AIPYH--PINPAPDVTAAAGGPPVROPAPLISAWPDQAQRPAVASRPPTTACAAP	109
	:	: : : : : : : : : : : : : : : : :	
	:	: : : : : : : : : : : : : : : : :	
DB	1867	ALPALRTATTTATSVTAI-----PSSSLGAN-----IFLSITTTTAT	1907
	:	: : : : : : : : : : : : : : : : :	
QY	110	LTVAPADTPPPDVDSRGAILRQYNLTSPFTSSVATGTNLVYAAPLSPLLPQQG	169
	:	: : : : : : : : : : : : : : : : :	
DB	1908	MSTATPS-SIPE-----TVHTSTVLTTATTTCGSVAISS-----TPG	1946
	:	: : : : : : : : : : : : : : : : :	

QY	170	T - NTHLMATEASNYALYPVAFATIPYPLPVNAGGYAISIISWPOTTTPT-----	220
DB	1947	TAHTTKVPDTTTCGT-----ATPSSSPGIALTPPWISTTTTTPTGSTVT	1993
QY	221	--SVMSLITSTVPHVQHGIASEVIPSEPIHPNQGWFSVTSVAHFHATSGLVM.	278
DB	1944	PSSPGTHATLVLTITTTTVAGSMATPS-----STSGSPSLTITATTIT	2043
QY	279	CIHGSILN-----SYNTPTYTGALG-----LLDPFALFEF	308
DB	2044	AI-GSINPSSFGITHPEVLITATPAATSSVTP-SSALGCTIHPWPVNITATHC	2101

QY	309	ENLTGNTNPVSPYSTAPHPPLPCGACGTAEIT	-----TTAAAT-----	447
DB	2102	KSLPSSPHVPTAWISAT	-----SCIGCTTHITFPSTCTSHTTAAATGCTGCTSTPAI	SS 2156
QY	348	-----RPMKULYIYSINGVETGEGCIAITLNI	-----AULTLGGLEPTE-----	LSSAG 192
DB	2157	PHPSRRTTESPPSPGTTTTGHT	EGTSRTATATPSKTRTSTILLPSSPTNAPITTVVTTG	2215
QY	343	QQ-----LPYSRPPVSGANGETPVKLYISVE	-----NAQDKGI-----	425
DB	2216	CFEPAQAWSEWIDVSYIMQPPSC	CHFDYUYNIPAAAGAGVCEQPLILEFACACPCVPIRF	2274

QY	426	---AIPHIDILG---	---PSKVVLDVDMQHQQP---	---TUSPAP	458
DB	2332	---GAAVVFSTFGLVTPRQVQKMKNYRIVPCCNYG	---HGTSTATSTATSTST---	2332	
QY	459	SKPSFVKANLVMLWSLHAARVD---	---QSLVGSSGLGVVYSUVLVNNVATGAQAV	510	
DB	2343	---ETTWILTK---	---LITATITESTATSTPSTGCPACTPHYSTATIPV	TSKAT	2385
QY	511	APSLDWIKVLDGMPISGLQVYSKIPFLVPLGKLSFWEACTKAGVYNYNNTASQIL		570	

Ddb 2386 PESSNGTATATAPKRSSTATTPTATSPALPSSSAGTTW-TRLSQTTTPMAINSTAIPSSI 2444

Qy 571 VENAACHRAVALSTTTSLGAGFV-LCAVAFLAHHSALALELGLHFFAAH 622
| | | | |
Db 2445 PEIVHTSVLTITATTTCATG-- SVATPSST-----PGIAHI 2479

RESULT 7

S52863 DNA binding protein R kappa B human
C-Species: Homo sapiens (man)
C.Date: 08 May 1995 sequence_revision 21 Jun 1995 #text_change 01-Dec-2000
C.Accession: S52863, A45580
R.Rieters, A.; Roumestier, T.; Scheidreit, C.
submitted to the EMBL Data Library, August 1994
A:Reference number: S52863
A:Accession: S52863
A.Synonyms: Preliminary
A.Molecule type: mRNA
A.Restrictions: 1 1224 -BRI-
A.Cross-references: EMBL:X80878; NID:a695578; PID:a695579
F.Abas, B.S., Le-Gat, F.Y., Haeley, E.W., Nobel, G.J.
New Biol. 3, 1063-1074, 1991
A.Title: Cloning of R kappa B, a novel DNA binding protein that recognizes the inter-
A:Reference number: A45580; MUID:92135142; PMID:1777480
A:Accession: A45580
A.Molecule type: mRNA
A.Restrictions: 314 1252 -ENSEPELLAF' 1264 'GGPLI' -ACA-
A.Cross references: DBJ:D8191, CR:S79520, MID:a476273, PID:a476274
A.Note: Sequence extracted from NEPI backbone (NEPI.N.79520)
C.Keywords: DNA binding; transcription factor

```
Query Match      4.08; Score 135.5; DB 2; Length 1324;
Best Local Similarity 19.0%, find: No 2;
Matches 131; Conservative .45; Mismatches 242; Indels 241; Gaps 29;
```

CY 6P AAGACGPEVPGAPP-----TGSAMFQDGPDAVASPRPT-----IA 105
||| || | | | | | | | | |
Db 6Q AAATFAKAECKPPKPFSSVSSEKFSKTSKTSLSMGLSISMSRRTVEVFII 752
||||| || | | | | | | | | |

CY 1GG GAAPL IAAVAHHLPPPVPSLVSGALIEPKYNLSIQLISSVALINVLAAARSEL-- 164
||||| || | | | | | | | | |
db 7C3 PALLAIPRP PVASANKSC-----ITGVSHAFSPGSMVNEEMPMIG- 708

QY	165	FLGDGINTIMATED:SNVQYR VAKATIKRYELVINAVRYALSI----	209
DB	799	TMLSPASSQAPSSQAAARVUSISNAGISQVWVAQPSI---PAPQDSNRGPAQLPQM	855
QY	210	SNWPTT TITFTVTRRSLTIDWRILVQVSNASHLY	245
DB	856	PAGEIQKVPATATQKVPQTVMTAVIVRAQTAAVQPPQDQICITVLSIPATASVQ	915
QY	246	TPSERLHYRQGRSVETISVAEEATSGILVMIITHGSLVNSYTPITPTALGI-----	299
DB	916	ALQVETTSAPQALAAVTRN-----VQVAVV-----VQVQV	961

[illegible]

QY	462	FSTFKNIVW----	SITAFVIG--SYGSSIEGVYSNVIL-----VN	582
	:	:	:	:
DB	1110	AATIRVGLAVPEPRADITIVAHAKCASVASNSVTIHS	AVSSEMNRAVKRTVA	1198
	:	:	:	:
OV	504	VATCAAGAAAGSIOW-----PKVLEWDPTSTTGASKTFEV	538	

Db 589 GSUTITAT-----DTTFTVSGSGPPLMDAALTAHAPGLTAP 639
 QY 402 FALEFFPNTPTNTFVSYSSTAP-----HPLPPGA----- 336
 Db 631 -----TLLTGTHTDLSHATTAAQPLPHGIMTTAGCHLTAYGTHTIQIPAPILD 682
 QY 336 -----DGAELTTTAAATPKMDLYFTSTNVGIGRGIAITIPNIADTLTGIPPEH 387
 Db 683 NTGSIATNTLDMHTAALENTGTHSTAT-----PNPFTTCTTINAGHLELNGPIL 739
 QY 488 I-----SSAGGOLFYSRPVV-----SANCEPTVKLYTSVNAQDQGIAPH 429
 Db 740 TIGTWTNTGSHQITGATHAITHATNPNGIHTATGPDLDITGILNN--GUNGELTTT 797
 QY 430 D-IDGERSVVOQYNQH-FQDRPTPSAP-----SPFESVIPANTVIM 473
 Db 798 DALTLTAASLI-----NORCTTAAAGPAHETLTGILNTAGSLQIAHILWLISASLINS 852
 QY 474 -SLTAAEYDOSTYSS-----TGPV-----YVSDSVTL 500
 Db 853 GTLASQTLITDQTHLIDNTGPMGTITGMLILHSPLNTAGLQIAAALTIDTAATL 912
 QY 501 VNVATGACAVARSIDKVTIDGRPLSTQQYSKTPEVPLRGKISFWAGCTKACYPVN 560
 Db 913 INRDGGALLAATLDTALTLDNRG-GTIDSGTAT-----HLRTTLDNTTAGHSS 963
 QY 561 YNTTASQQLLVENNAAG 576
 Db 964 SGTLOIDGTLTNTGG 979

RESULT 10

B82519

hemagglutinin-like secreted protein XF2775 [imported] Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20 Aug 2000

C:Accession: B82519

Ranonymous: The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: B82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: B82519

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1 3455 SSIM

A:Cross references: GR A600487; GH AE001849; NID 93108063; FIIN AAF65601; ASPHA 30001

A:Experimental source: strain 9a5c

R:Simpsom A I G : Rainach, F C : Arruda, P : Abreu, F A : Accencio, M : Alvares, R : B. Briones, M E S : Bueno, M F E : Camargo, A A : Canha, L E A : Carriao, D M : Gallet, R as Neto, E : Ducloux, C : El-Jabary, H : Fariouani, A P : Ferreira, A J S.

submitted 30 Dec 2000, June 2003

A:Authors: Ferreira, V C A : Fritov, J A : Fraga, J S : Franco, S C : Franco, M C : Fraga, J D : Junqueira, M L : Kemper, E L : Kitajima, J P : Krueger, J E : Kurumaji, E E : Lago chado, M A : Madeira, A B N : Madeira, H M F : Marino, C L : Marques, M V : Martins, E A:Authors: Martins, E M F : Matsukuma, A Y : Meuck, C F M : Miracca, E C : Miyaki, C Y : F. G. : Nunes, L R : Oliveira, M A : de Oliveira, M C : de Oliveira, R C : Palmieri, D A : Rodrigues, V : Rosa, A I de M : de Rosa, V E : de Sa, P G : Santelli, R V : Sawasak A:Authors: da Silva, A C P : da Silva, F R : da Silva, A M : Silva Jr., W A : da Silva, M : Isibako, M H : Valada, H : Van Sloys, M A : Vitoriski, Almeida, S : Vitoriski, A

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF2775

Query Match 4.0% Score 135.5 DB 2 Length 3455

Best Local Similarity 21.9% Pred. No. 8;

Matches 135; Conservativ 64; Mismatches 216; Indels 201; Gaps 27;

QY 97 ASREPTTAGAELTAVAHDTF-----FVP-----QVSGGAIL 122

Db 420 ASPADPT-----PVASVPPATPPPTGMPIPSPQATAPVLAAGKLTLENDLNPOCHI 483

QY 122 PPOVNFCTSPSESAVACGNNLVYAALDSPLERLQCTNTH-----IMAFEA-----SN 181
 Db 424 TAGGATVAITLIDNPT-----AALNPLTLCGGLPNAAGCTTTATFATPTPTNN 537
 QY 182 YACVAVARATIPPEPIVPAAGVYASISFWQTHITPVSVMNSLITDVRHIVAPSA 241
 Db 538 AACQI PVNFTIDTAQHI NNASHLITP--POSAPHTIADLNTSSTE-----ASA 588
 QY 242 SEVATFDELHYRGWGSVLTGVALDLATSELVWHTHELVRYVNTPTVGAAGLID 601
 Db 589 GSULTAATL-----DTTGTGSGSGPPLMDAALTAHAPGLTAP 639
 QY 302 FALEFFPNTPTNTFVSYSSTAP-----HPLPPGA----- 336
 Db 631 -----TLLTGTHTDLSHATTAAQPLPHGIMTTAGCHLTAYGTHTIQIPAPILD 682
 QY 336 -----DGAELTTTAAATPKMDLYFTSTNVGIGRGIAITIPNIADTLTGIPPEH 387
 Db 683 NTGSIATNTLDMHTAALENTGTHSTAT-----PNPFTTCTTINAGHLELNGPIL 739
 QY 488 I-----SSAGGOLFYSRPVV-----SANCEPTVKLYTSVNAQDQGIAPH 429
 Db 740 TIGTWTNTGSHQITGATHAITHATNPNGIHTATGPDLDITGILNN--GUNGELTTT 797
 QY 430 D-IDGERSVVOQYNQH-FQDRPTPSAP-----SPFESVIPANTVIM 473
 Db 798 DALTLTAASLI-----NORCTTAAAGPAHETLTGILNTAGSLQIAHILWLISASLINS 852
 QY 474 -SLTAAEYDOSTYSS-----TGPV-----YVSDSVTL 500
 Db 853 GTLASQTLITDQTHLIDNTGPMGTITGMLILHSPLNTAGLQIAAALTIDTAATL 912
 QY 501 VNVATGACAVARSIDKVTIDGRPLSTQQYSKTPEVPLRGKISFWAGCTKACYPVN 560
 Db 913 INRDGGALLAATLDTALTLDNRG-GTIDSGTAT-----HLRTTLDNTTAGHSS 963
 QY 561 YNTTASQQLLVENNAAG 576
 Db 964 SGTLOIDGTLTNTGG 979

RESULT 11

T21460

hypothetical protein XK945.10 Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T21460

A:Status: preliminary

A:Cross references: IMB148052; FIIN CAA8849.1; GSPDB GNS00020; GESP EK945.10

submitted to the EMBL Data Library, March 1995

A:Reference number: 219425

A:Accession: T21460

A:Molecule type: DNA

A:Residues: 1 796 WIL-

A:Cross references: IMB148052; FIIN CAA8849.1; GSPDB GNS00020; GESP EK945.10

A:Experimental source: clone F27E5

C:Genetics:

A:Gene: GESP:EK945.10

A:Map position: 2

A:Contents: 523 124 3 1647 124 3 273 273 273 273 273 273

Query Match

3.9% Score 139; DB 2; Length 796;

Best Local Similarity 20.6% Pred. No. 1.4;

Matches 121; Conservativ 61; Mismatches 228; Indels 176; Gaps 24;

QY 63 APTVTAATACGPPVCPASPHGSANPCCAGANAVASPPPPASADAPITAPAPHIDV 122

Db 242 ARGEAAGASPHLEMEELAGLNLTALLESISTATTTTITVTTTSTIV 296

QY 123 TATVNGAALERYRLETLTSSVATGTMVLYAATSTELTLCGNTNTHMAHRYN 182

Db 297 P-----TSTCTVTAMSTSS-----TCTSTTSTSTSTSTSTSTSTST 339

```

QY 163 AGPPVAKATIKYR-----LVNAVQGY-----AISISW---212
DB 339 STIGSSSTITSSSSSTISSTISPTTTPETITSLSPENACVSDITETITTTT 368
QY 213-----PQTTTPTSDMS-----ITSDVRILVQGIASELVIPS---248
DB 399 LISTITREPTSTTTTETWSTESSVTTTETPTTLTSTASTSTETPSITVTSPISPV 458
QY 249-ERLIYRQGWKSVTSVAFEPATSGVLMGLHGSILVNSYNTPTGALGLDFALELE 307
DB 459 TSVISSSSSSSVIIPSTIESIST-----PSSIVITSTAPSTISITG-----PS 504
QY 608 ENITPGNT-----NTRVSVYSSTARRLRGADGTAELTTTAAATRFMKDLYF-----355
DB 505 SSSSTPSSASSSSSVSIASSIUSST--TGQOSTTKTSEITSSXTNDFEVEPKATT 561
QY 636-----TSIN-----GVDEIGIGIAL-----ILFNLADEL-----LGG-LPT 485
DB 562 TPYSTSVNLHNSHGHIDYGLSTSTSTSSSVVSTTKMGATKTSVSMPLGGIYDA 421
QY 606 ELISSAGGLVFSRPSVVSANGPTVKLYTSVENAQURKGLALPHDIDGESKVVIOUDYN 445
DB 622 STFVYDQWTF RAIMTIDRK---KVYTYAN-----VYIQYSS 657
QY 446 QH EQDRI-----PSPAPSRPESVLKANIWLMLSLTAARYDQSTVGSSTGPPVVS---495
DB 658 TTIESPSSISAVASSNPSSTPSSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 717
QY 496 --DSVLNVNVALGAGAVAKSLIKVILDKRPLSTIOYSKTFVL 539
DB 718 SPSTHDSSTITSTSTSSPSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 760

RESULT 12
S675704
Hsp61 protein precursor: Yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YDR420W
C:Species: Saccharomyces cerevisiae
C:Date: 23 Apr 1996, #sequence_revision: 96, 9p 1996, #text_change: 23 Mar 2001
C:Accession: S67570; A54882
Ref:Kobayashi, F.S.
submitted to the EMBL data library, August 1995
A:Description: The sequence of S. cerevisiae lambda 3641 and cosmids 9461, 9467, and 9468
A:Reference number: S67555
A:Accession: S67570
A:Molecule type: DNA
A:Residues: 1 1802 aa
A:Cross references: EMBL:014007; NID:0927695; PDB:AAB64857.1; PDB:0927691; MIPS:YDR420W
R:Kasahara, S.; Yamada, R.; Miyazaki, Y.; Shiratori, Y.; Miyamoto, C.; Yabe, T.; Nakajima, T.
J. Biochem. 176, 1488-1499, 1994
A:Title: Cloning of the Saccharomyces cerevisiae gene whose overexpression overcomes the
A:Reference number: A54882; M01194156857; EMBL:014151
A:Accession: A54882
A:Molecule type: DNA
A:Residues: 1 581, A' 583-593, A' 595-1802 aa
A:Cross references: EMBL:S69101; NID:0545659; PDB:AAB30051.1; PDB:0545660
A:Experimental source: YRN295
A:Notes: Sequence extracted from NBI backbone (NCBIN:144410, NCBI:P144411)
C:Date of res:
A:Accession: S67571
A:Molecule type: DNA
A:Residues: 1 583 aa
A:Cross references: S67570; S67571; MIPS:YDR420W
A:Map position: 49
C:Keywords: calcium binding; glycoprotein; transmembrane protein
F:1 21/26aa: signal sequence status predicted: SIG
F:22 1002/1002aa: msp protein status predicted: MAF
F:1493 1108/1008aa: transmembrane status predicted: TM
F:1645 1656/1008aa: calcium binding status predicted: CAL
Query Match 4.9% Score 143; DB 2; Length 1802;
Best Local Similarity 19.0% Pred. No. 4.5;
Matches 140; Conservative 193; Mismatches 264; Indels 188; Gaps 40;

```

```

QY 48 SUPALPYHTNPENAPLPIVIAAAAGHVRVRCAPR LGSARKQAGRIAVASRRKGLA 105
DB 705 SAPAVNSSTYISSAP AATSSYISFSAFVAVSYIYSSSAPAAATSSV 155 709
QY 136 GAALTAVALAHDTVPVWVERKATLERATNRLTELEELVATLRLVLYAAHLSLELP 245
DB 760 PSAIV AVSSYIYSSSALVWLSSTSSSYDVIWVSST FAATSSSYLP 807
QY 166 LAGTINIDIMALEASNY AGRVARALITRYLIVNAV 3541 207
DB 808 SPASVAMSSSTSSSPVLYVSSSSASSTIAIYFSSSTSSSTSSSTSSSTSSSTSSST 867
QY 208 SLSIWFQTTT-----PTSVMMNSTLSTHARILVAGNCHACHIVDSEILHVNQGWKSVLS 264
DB 868 EFSSEKSTSTLAPTOHSILSKLID FLQHSMAUCSTSVQ 2515 914
QY 295 QVAHEALSHVME FTRSSVRYVYRITFY GASHHLLALHELEN 1 311
DB 915 STLNIEHSSALSVNPSASNLVPSITLSTLQASHLSRKNKATPSLQASHLSRKNY 974
QY 912 TPGN'NR VGRY SSTADHPLRFAHGLABELLTAALIKMKGLY 131 360
DB 975 HAKNHLHSRSTVGRFAYLSAAKHNKLSATLFTETLSSSHAYLHLSHRS 1364
QY 361 WREGRG---ALHTFNLAHLDGRL PLE TSSAGGLVTSRVVAVANJE 407
DB 1036 VTGSHBHVLSGFALTYSGTGLRFLLELTHAKAVAGLQKRLHJLGLIGANSQ 1094
QY 408 -----PIVKLSTSVNAQ QPKHALER 429
DB 1095 SPKSRKSVAGSSSHSGSGLRFRVYVSSSGSEFVKTDRGRDFGLLFLK 1154
QY 440 -----DLDGSKVVLQYDQHQE LKPTGSPALSRFSVLKAVIMLS 474
DB 1155 SSSTSSSYSKSTKLSIGALHAGVQASINSVFLAAAVYLYFP SENGAKL 1212
QY 475 LTAALYFSGTYNSGHVYVGRVLYVVAALAAVAGGLWKV 1193 523
DB 1213 TATVLSGELRTAENGLGLD NALLAAVAFEL NBLHLLHETGAALRYVVLV 1271
QY 524 RNSGTYNSKTYVTHLPRKLELWALHLEKAVYVYR LARGLAVRMAAARVVAIS 1603
DB 1272 NPLSSAQ---TFNPLHLVKK YFENQSSSE LNSLSE 157 1607
QY 584 YTTSLGAG---FWSISAVAVL 601
DB 1408 FILSVPSSSSTILSRKSTSSLSV 1407

RESULT 13
S67571
Hsp61 protein Ydr420w: yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein b2726
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996, #sequence_revision: 12, Jul 1996, #text_change: 19 Apr 2003
C:Accession: S67571
R:Paulin, L.; Søren, A.M.; Laamanen, P.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67560
A:Accession: S67571
A:Molecule type: DNA
A:Residues: 1 583 aa
A:Cross references: EMBL:Z74087; NID:01431621; PDB:0275901; PDB:01431621; MIPS:YDR420W
A:Experimental source: YRN295
C:Date of res:
A:Accession: S67571
A:Molecule type: DNA
A:Residues: 1 583 aa
A:Cross references: S67570; S67571; MIPS:YDR420W
A:Map position: 49
C:Keywords: p18 submaxillary protein
Query Match 4.9% Score 142.9; DB 2; Length 980;
Best Local Similarity 19.0% Pred. No. 4.5;
Matches 112; Conservative 160; Mismatches 264; Indels 167; Gaps 18;

```



```

QY 181 NYAGYVPVAPATIPYPPPIVNVAVGGYALISSEWPGTITITPTISVFMNSITSTVPIIYVPGPI 240
DB 181 NYAGYVPVAPATIPYPPPIVNVAVGGYALISSEWPGTITITPTISVFMNSITSTVPIIYVPGPI 240
QY 241 ASEIVIPSEIRHYRNOGWRVSVETSGVAFFRATSI VMLC IHGSLVNSVNTPTVTSALGILL 300
DB 241 ASEIVIPSEIRHYRNOGWRVSVETSGVAFFRATSI VMLC IHGSLVNSVNTPTVTSALGILL 300
QY 301 DPALELEFRNLTPGNTNTRVRSYSSSTARHLRGAGDTAELTTTATATRMKKOLYFTSTNG 360
DB 301 DPALELEFRNLTPGNTNTRVRSYSSSTARHLRGAGDTAELTTTATATRMKKOLYFTSTNG 360
QY 361 VGEIGRGIALTFENLALITLGGTPTFISSAGGGLFYSPVSVANSSEPIVKLYISVENAQ 420
DB 361 VGEIGRGIALTFENLALITLGGTPTFISSAGGGLFYSPVSVANSSEPIVKLYISVENAQ 420
QY 421 QDKGIATPHDIDIGESRVVIGQYDNQHEQDRPTSPAPSPFPFVSLFANDVLMLSLTAABY 480
DB 421 QDKGIATPHDIDIGESRVVIGQYDNQHEQDRPTSPAPSPFPFVSLFANDVLMLSLTAABY 480
QY 481 DQSTYGSSTGPGVVSUSVTLVNVATGAGAVAPSLDWTKTVDGKPISTTQOYSKTFPVLP 540
DB 481 DQSTYGSSTGPGVVSUSVTLVNVATGAGAVAPSLDWTKTVDGKPISTTQOYSKTFPVLP 540
QY 541 IPGKISFWEAGTTKAGYPYNYNTTASQGLIVENAGHPVAISYTTSLGAGPVSISAVAV 600
DB 541 IPGKISFWEAGTTKAGYPYNYNTTASQGLIVENAGHPVAISYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALALIEDTLDYPARAHITFDGPPPCRPPIGIGTAPGTSVAHLQPKMKVSKTREL 660
DB 601 LAPHSALALIEDTLDYPARAHITFDGPPPCRPPIGIGTAPGTSVAHLQPKMKVSKTREL 660

RESULT 2
VST2_HEVPA
ID VST2_HEVPA STANDARD: PPT: 660 AA
AC P33426;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Pakistan) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses
OX NCBI_TaxID=33774;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.O., Reyes G.R., Tsareva T.S., Lecters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RI "Characterization of a prototype strain of hepatitis E virus.";
RL Proc Natl Acad Sci U S A 89:550-563(1992)
CC 1-1 FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED PNA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announcements
CC or send an email to license@isb-sib.ch.)
CC
DR EMBL: M80581; AAA45727.1; -
DR InterPro: IPR004261; SP2
DR Pfam: PF03014; SP2; 1.
KW Signal.
FT SIGNAL. 1 22 BY SIMILARITY.
FT CHAIN. 23 660 STRUCTURAL PROTEIN 2.
SQ SP2:FNPF 660 AA: 70080 MW: 80856.530 PM515F03 APC64;

```

```

Query Match 99.5%; Score 3391; DB 1; Length 660;
Best Local Similarity 99.4%; Prob. 2.2e-206;
Matches 656; Conservative 1; Mismatches 5; Indels 6; Gaps 0;
QY 1 MRPPIIIIIIMPELPMIPAPPPQPPSPRRRRSGSGSGGFWDRVDSQFFPAIYIHPIN 60
DB 1 MRPPIIIIIIMPELPMIPAPPPQPPSPRRRRSGSGSGGFWDRVDSQFFPAIYIHPIN 60
QY 61 PFADVTAAACAGCPVPQPPAPPIIGSAAWPGCAQPPAVASHPRPPIIAAAILTAVAAHDTF 120
DB 61 PFADVTAAACAGCPVPQPPAPPIIGSAAWPGCAQPPAAASRPPTTAAAPPIAVAAHDTF 120
QY 121 PVQVDPSEKATIRPQYNI STSPITSSVATGNI VI VAAPI SDI PTQNTNTHIMAFAS 180
DB 121 PVQVDPSEKATIRPQYNI STSPITSSVATGNI VI VAAPI SDI PTQNTNTHIMAFAS 180
QY 181 NYAGYVPVAPATIPYPPPIVNVAVGGYALISSEWPGTITITPTISVFMNSITSTVPIIYVPGPI 240
DB 181 NYAGYVPVAPATIPYPPPIVNVAVGGYALISSEWPGTITITPTISVFMNSITSTVPIIYVPGPI 240
QY 241 ASEIVIPSEIRHYRNOGWRVSVETSGVAFFRATSI VMLC IHGSLVNSVNTPTVTSALGILL 300
DB 241 ASEIVIPSEIRHYRNOGWRVSVETSGVAFFRATSI VMLC IHGSLVNSVNTPTVTSALGILL 300
QY 301 DPALELEFRNLTPGNTNTRVRSYSSSTARHLRGAGDTAELTTTATATRMKKOLYFTSTNG 360
DB 301 DPALELEFRNLTPGNTNTRVRSYSSSTARHLRGAGDTAELTTTATATRMKKOLYFTSTNG 360
QY 361 VGEIGRGIALTFENLALITLGGTPTFISSAGGGLFYSPVSVANSSEPIVKLYISVENAQ 420
DB 361 VGEIGRGIALTFENLALITLGGTPTFISSAGGGLFYSPVSVANSSEPIVKLYISVENAQ 420
QY 421 QDKGIATPHDIDIGESRVVIGQYDNQHEQDRPTSPAPSPFPFVSLFANDVLMLSLTAABY 480
DB 421 QDKGIATPHDIDIGESRVVIGQYDNQHEQDRPTSPAPSPFPFVSLFANDVLMLSLTAABY 480
QY 481 DQSTYGSSTGPGVVSUSVTLVNVATGAGAVAPSLDWTKTVDGKPISTTQOYSKTFPVLP 540
DB 481 DQSTYGSSTGPGVVSUSVTLVNVATGAGAVAPSLDWTKTVDGKPISTTQOYSKTFPVLP 540
QY 541 IPGKISFWEAGTTKAGYPYNYNTTASQGLIVENAGHPVAISYTTSLGAGPVSISAVAV 600
DB 541 IPGKISFWEAGTTKAGYPYNYNTTASQGLIVENAGHPVAISYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALALIEDTLDYPARAHITFDGPPPCRPPIGIGTAPGTSVAHLQPKMKVSKTREL 660
DB 601 LAPHSALALIEDTLDYPARAHITFDGPPPCRPPIGIGTAPGTSVAHLQPKMKVSKTREL 660

RESULT 3
VST2_HEVMY
ID VST2_HEVMY STANDARD: PPT: 660 AA.
AC Q04611;
DT 01-OCT-1994 (Rel. 27, Created)
DT 01-OCT-1994 (Rel. 27, Last sequence update)
DT 01-NOV-1995 (Rel. 42, Last annotation update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Myanmar) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=41769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9322757; PubMed 9470471;
RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
RA Ekihiisa T., Wain K.;
RI "Sequence and gene structure of the hepatitis E virus isolated from
RI Myanmar.";
RL Virus Genes 7:95-109(1994).
CC 1-1 FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED PNA.
CC

```

cc this SWISS prot entry is copyrighted. It is produced through a collaboration
cc between the Swiss Institute of Bioinformatics and the EMBL outstation.
cc the European Bioinformatics Institute. There are no restrictions on its
cc use by non-profit institutions as long as its content is in no way
cc modified and this statement is not removed. Usage by and for commercial
cc entities requires a license agreement (see <http://www.ebi.ac.uk/infocentre>
cc or send an email to license@ebi.ac.uk).

EMBL: D19450; GenBank: U001174.1;
DR InterPro: IPR004261; SPZ;
DR Pfam: PF04014; SPZ; 1;
KW SIGNAL;
FT SIGNAL 1 22 BY SIMILARITY;
FT CHAIN 23 660 STRUCTURAL PROTEIN 2;
SQ SEQUENCE 660 AA; 70998 MW; 9A82A4EA155C552 CRC64;

Query Match 98.8%; Score 3367; DB 1; Length 660;
Best Local Similarity 98.8%; Pred No. 1e-204;
Matches 654; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 MRPRILLLLMLPMLPAPPGQISGRFRGRSGSGGFWGPRVLSQPFALPYTHPTN 60
DB 1 MRPRILLLLMLPMLPAPPGQISGRFRGRSGSGGFWGPRVLSQPFALPYTHPTN 60
QY 61 PFAPVTAAGAGPRVROPAPRPLGSAMRQAGRPVAVASRRRPITAGAAPLTAVAPAHDP 120
DB 61 PFAPVTAAGAGPRVROPAPRPLGSAMRQAGRPVAVASRRRPITAGAAPLTAVAPAHDP 120
QY 121 FVPVDSKCALTRQYNLSISPLSSVATGNIIVLYAAPLSPLPLPQGNTHIMAEAS 180
DB 121 FVPVDSKCALTRQYNLSISPLSSVATGNIIVLYAAPLSPLPLPQGNTHIMAEAS 180
QY 122 FVTLGKRAALTRQYRSTSPFTSVATGNIIVLYAALESLLLSQGNTHIMAEAS 180
DB 122 FVTLGKRAALTRQYRSTSPFTSVATGNIIVLYAALESLLLSQGNTHIMAEAS 180
QY 181 NYAQYRVKATIRYRPLVPNAVGYAISSEFWPTITPTPSVMNSITSDVRLVQPGI 240
DB 181 NYAQYRVKATIRYRPLVPNAVGYAISSEFWPTITPTPSVMNSITSDVRLVQPGI 240
QY 241 ASELVIPSERHYRNPQWRVSVETSGVAFPAATSLVMLATGSLVNSYNTPTVGAIGLL 300
DB 241 ASELVIPSERHYRNPQWRVSVETSGVAFPAATSLVMLATGSLVNSYNTPTVGAIGLL 300
QY 301 DFALLEPRLNLTPTNTIRVSRYSSTAPHLRPGADGTAFITTAATRFMKDLFTSTNG 360
DB 301 DFALLEPRLNLTPTNTIRVSRYSSTAPHLRPGADGTAFITTAATRFMKDLFTSTNG 360
QY 361 VHEIQRGLALFLNLATLLAGLPTELISAGGQLFYSRPVVSANGRPVTKLYTSVENAQ 420
DB 361 VHEIQRGLALFLNLATLLAGLPTELISAGGQLFYSRPVVSANGRPVTKLYTSVENAQ 420
QY 421 QIKGLAIPHDILGDSKVVIGQYUNGHPQDRPTSPAPSRPFSLRANGVLSLTAAY 480
DB 421 QIKGLAIPHDILGDSKVVIGQYUNGHPQDRPTSPAPSRPFSLRANGVLSLTAAY 480
QY 481 DQSTVGSSTGVVYSDSVTLNVNVAIGAVARSIDWTKVTLDRPLSTIQQYSKTFEVL 540
DB 481 DQSTVGSSTGVVYSDSVTLNVNVAIGAVARSIDWTKVTLDRPLSTIQQYSKTFEVL 540
QY 541 LRKGLSEWAGITKAGHPYNTNTIASDQLLVENAGRHVALSYITTSIAGPVSISAVAV 600
DB 541 LRKGLSEWAGITKAGHPYNTNTIASDQLLVENAGRHVALSYITTSIAGPVSISAVAV 600
QY 601 LAPHSALALLEDLIDYFACAHFDLDRCPHRLPIALQCAFDQSVIAELQRLKMKVGKTR 660
DB 601 LAPHSALALLEDLIDYFACAHFDLDRCPHRLPIALQCAFDQSVIAELQRLKMKVGKTR 660

```

RESULT 4
VSL2 BEVME
AD VSL2 BEVME STANDARD; PRT; 65% AA
AC 60560;
DI 01 OCT 1993 (601, 27, Created)
DI 01 OCT 1993 (601, 27, Last sequence update)
DI 01 FEB 1994 (601, 28, Last annotation update)
DE Structural protein 2 precursor.

cc Hepatitis E virus (strain Mexico (HUV)
cc Virus; ssRNA positive strand viruses; no DNA stage;
cc Hepatitis E-like viruses;
cc NCBI_TaxID 41768;
RN 11;
SI 1; ID 1; 2; 3; 4;
RX MEDLINE 94079657; PubMed 1448913;
RA Huang C C, Nguyen D, Fernandez J, Yun K Y, Fiy K E,
RI Bradley D W, Tan A W, Reyes G P,
RT Molecular cloning and sequencing of the Mexican isolate of hepatitis
E virus (HEV).
RL Virology 191:560-568(1992).
CC 1 FUNCTION: CONTAINS A HIGH BASIC AMINO-ACID-CONTENT GREGG-119;
CC THAT IT MAY BE INVOLVED IN THE INACTIVATION OF THE GREGG-19A
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC

cc THIS SWISS-prot entry is copyrighted. It is produced through a collaboration
cc between the Swiss Institute of Bioinformatics and the EMBL outstation.
cc the European Bioinformatics Institute. There are no restrictions on its
cc use by non-profit institutions as long as its content is in no way
cc modified and this statement is not removed. Usage by and for commercial
cc entities requires a license agreement (see <http://www.ebi.ac.uk/infocentre>
cc or send an email to license@ebi.ac.uk).

EMBL: M74506; AAA5732.1;
DR PIR: B44212; B44212;
DR InterPro: IPR004261; SPZ;
DR Pfam: PF04014; SPZ; 1;
KW SIGNAL;
FT SIGNAL 1 22 BY SIMILARITY;
FT CHAIN 23 659 STRUCTURAL PROTEIN 2;
SQ SEQUENCE 659 AA; 70640 MW; 6175E7E9F9E9BEC CRC64;

Query Match 94.7%; Score 4194.5; DB 1; Length 659;
Best Local Similarity 94.0%; Pred No. 7.5e-194;
Matches 614; Conservative 21; Mismatches 24; Indels 1; Gaps 1;

```

QY 1 MRPRILLLLMLPMLPAPPGQISGRFRGRSGSGGFWGPRVLSQPFALPYTHPTN 60
DB 1 MRPRILLLLMLPMLPAPPGQISGRFRGRSGSGGFWGPRVLSQPFALPYTHPTN 60
QY 61 PFAPVTAAGAGPRVROPAPRPLGSAMRQAGRPVAVASRRRPITAGAAPLTAVAPAHDP 120
DB 61 PFAPVTAAGAGPRVROPAPRPLGSAMRQAGRPVAVASRRRPITAGAAPLTAVAPAHDP 120
QY 121 FVPVDSKCALTRQYNLSISPLSSVATGNIIVLYAAPLSPLPLPQGNTHIMAEAS 180
DB 121 FVPVDSKCALTRQYNLSISPLSSVATGNIIVLYAAPLSPLPLPQGNTHIMAEAS 180
QY 181 NYAQYRVKATIRYRPLVPNAVGYAISSEFWPTITPTPSVMNSITSDVRLVQPGI 240
DB 181 NYAQYRVKATIRYRPLVPNAVGYAISSEFWPTITPTPSVMNSITSDVRLVQPGI 240
QY 241 ASELVIPSERHYRNPQWRVSVETSGVAFPAATSLVMLATGSLVNSYNTPTVGAIGLL 300
DB 241 ASELVIPSERHYRNPQWRVSVETSGVAFPAATSLVMLATGSLVNSYNTPTVGAIGLL 300
QY 301 DFALLEPRLNLTPTNTIRVSRYSSTAPHLRPGADGTAFITTAATRFMKDLFTSTNG 360
DB 301 DFALLEPRLNLTPTNTIRVSRYSSTAPHLRPGADGTAFITTAATRFMKDLFTSTNG 360
QY 361 VHEIQRGLALFLNLATLLAGLPTELISAGGQLFYSRPVVSANGRPVTKLYTSVENAQ 420
DB 361 VHEIQRGLALFLNLATLLAGLPTELISAGGQLFYSRPVVSANGRPVTKLYTSVENAQ 420
QY 421 QIKGLAIPHDILGDSKVVIGQYUNGHPQDRPTSPAPSRPFSLRANGVLSLTAAY 480
DB 421 QIKGLAIPHDILGDSKVVIGQYUNGHPQDRPTSPAPSRPFSLRANGVLSLTAAY 480
QY 481 DQSTVGSSTGVVYSDSVTLNVNVAIGAVARSIDWTKVTLDRPLSTIQQYSKTFEVL 540
DB 481 DQSTVGSSTGVVYSDSVTLNVNVAIGAVARSIDWTKVTLDRPLSTIQQYSKTFEVL 540
QY 541 LRKGLSEWAGITKAGHPYNTNTIASDQLLVENAGRHVALSYITTSIAGPVSISAVAV 600
DB 541 LRKGLSEWAGITKAGHPYNTNTIASDQLLVENAGRHVALSYITTSIAGPVSISAVAV 600
QY 601 LAPHSALALLEDLIDYFACAHFDLDRCPHRLPIALQCAFDQSVIAELQRLKMKVGKTR 660
DB 601 LAPHSALALLEDLIDYFACAHFDLDRCPHRLPIALQCAFDQSVIAELQRLKMKVGKTR 660

```

```

QY 541 LRGKLSFWFACTIAAGHPYNNYNIASQILVENAAGHPVAISTYTHSIAGSPVSIASAVV 400
DB 540 LRGKLSWEAGTTKAGYPYNNYTASQILVENAAGHPVAISTYTHSIAGSPVSIASAAV 500
QY 601 LAPHSAALLLEDLDYPARAHITDDPCPCPCRPPLGLGLOGCAFOSTVAFLOPLKMKVGGKTRFL 660
DB 600 LAPPASALLLETHFYNGAPAHITFUCPCPCPCPAIGLGGTAPQGSIVAFLOPLKMKVGGKTRPHI 650

RESULT 5
ID VST2_HVHRH STANDARD; PRT; 485 AA.
AC Q00270;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE Structural protein 2 (Fragment).
OS Hepatitis E virus (isolate phcus) (HEV)
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
CC Hepatitis E-like viruses.
CX NCBI_taxid=31766;
RN [1]
RP SEQUENCE FROM N A
RX MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Ogasawara S.,
RA Wang C.-K., Shikata T., Ichikawa M., Iikihisa T., Mizuno K.,
RA Wio K.M.;
RT "Hepatitis E virus: cDNA cloning and expression."
RL Microbiol Immunol 36:67-79(1992).
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annotation/
CC or send an email to license@sib-sib.ch).
CC
DR EMRL: p90274; RAAQ0910 1; -
DR InterPro: IPR004261; SP2.
DR Pfam: PF03014; SP2; 1;
FT NON_TER 1 485
FT NON_TER 485
SQ SEQUENCE 485 AA: 52-417 MW: 544509272.72747494 CACG64;
Query Match
Best Local Similarity 90.0% Prod No 6 ae 147;
Matches 480; Conservative 1; Mismatches 4; Indels 6; Gaps 6;
QY 133 RQYNLSTSPITSSVATGTHNVIYAAPISPLPLDQGTNTHIMATPASNYAYVVPAPATL 192
DB 1 RQYNLSTSPITSSVATGTHNVIYAAPISPLPLDQGTNTHIMATPASNYAYVVPAPATL 60
QY 193 RYPLVFNAGGVAYSTSEWQPTTITPISVMNNSIISLAVPLLVGNGTASELVVISEPLH 252
DB 61 RYPLVFNAGGVAYSTSEWQPTTITPISVMNNSIISLAVPLLVGNGTASELVVISEPLH 120
QY 253 YPNAGFVSFVTSQVAFEFATSGIVMLCIGFSLVNSYNTPTGAGLLDFALEFPNLT 312
DB 121 YPNAGFVSFVTSQVAFEFATSGIVMLCIGFSLVNSYNTPTGAGLLDFALEFPNLT 180
QY 313 PCNTNTVSPYSSTAPHPPIPGAGCTAPITTTAATPFPMQVLYFTSTNGVCEIGRGIALTL 372
DB 181 PCNTNTVSPYSSTAPHPPIPGAGCTAPITTTAATPFPMQVLYFTSTNGVCEIGRGIALTL 240
QY 373 PNAATLIGGIPFLTSSAGGCTFYSGFVVSARHTFTVFTSVTHAAGCFGLAIFHHT 430
DB 241 PNAATLIGGIPFLTSSAGGCTFYSGFVVSARHTFTVFTSVTHAAGCFGLAIFHHT 300
QY 433 LGFSPVVYQYDNGHCFQVHPPTSPAPSPFPFVLPANVLMVLSLTAEEVQSTYSGSTCFV 492
DB 301 LGFSPVVYQYDNGHCFQVHPPTSPAPSPFPFVLPANVLMVLSLTAEEVQSTYSGSTAPV 360

```

```

QY 493 VVSNVSVLVNVAAGAGAVAPASITWTVPVIGSAPPISTLGGVSPSTFVVLHAPKISWEFAT 552
DB 361 VVSNVSVLVNVAAGAGAVAPASITWTVPVIGSAPPISTLGGVSPSTFVVLHAPKISWEFAT 420
QY 553 TRAGYPYNNYTASDQLLVENAGHPVAISTYTHSIAGSPVSIASAVVAPHSALALLED 612
DB 421 TRAGYPYNNYTASDQLLVENAGHPVAISTYTHSIAGSPVSIASAVVAPHSALALLED 480
QY 613 TLDYP 617
DB 481 TLDYP 485

RESULT 6
ID EGT2_YEAST STANDARD; PRT; 1041 AA.
AC P42835;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 42, Last sequence update)
DE EGT2 protein precursor (Early GI transcript 2).
OS Saccharomyces cerevisiae (Baker's yeast).
CC Eukaryote, Fungi, Ascomycota, Saccharomycotina, Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
CX NCBI_taxid=4942;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=S2886 / FY1676; PubMed=7645347;
RA Mifflahi M., Ni-and J. M., Levesque R., Gallardin C.;
RT "Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV
RT identifies the RPS4, PAB1 and RPL1 loci. Live now open reading
RT frames."
RL Yeast 11:567-572(1995).
RN [2]
RP SEQUENCE FROM N.A. AND FUNCTION.
RX MEDLINE=96251274; PubMed=8668141;
RA Kraschke P., Hamath K., Schuster T.;
RT "EGT2 gene transcription is induced predominantly by Swi5 in early
RT G1."
RL Mol. Cell. Biol. 16:3264-3274(1996).
CC -1- FUNCTION: SEEMS TO BE INVOLVED IN THE PRECISE TIMING OF CELL
CC SEPARATION AFTER CYTOKINESIS. AN ELONGATION OF MITOTIC DAUGHTER
CC CELLS IS DELAYED. COULD EITHER BE AN INTERMEDIATE NECESSARY FOR GLYCANS-
CC DEGRADATION OF THE CELL WALL AT THE NEXT MITOTIC DIVISION OR
CC DAUGHTER CELLS. A REDUCED BY-PRODUCT OF THIS METABOLIC
CC STEP. EXCLUSIVELY EXPRESSED BETWEEN THE END OF MITOSIS AND EARLY
CC G1. INACTIVATED BEFORE CELLS PASS START.
CC
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annotation/
CC or send an email to license@sib-sib.ch).
CC
DR EMRL: 246259; CAA86371.1; -
DR EMRL: 271603; CAA96259.1; -
DR PIR: S554862; S55862.
DR SCN: S0905271; EGT2.
DR GO: G01009277; Cell wall (source Fungi); IDA.
DR GO: G01016229; Proliferation; IMP.
KW Glycoprotein; Repeat; Signal; Cell cycle; Cell division; Mitosis.
FT CHAIN 1 1041
FT DOMAIN 21 1041 EGT2 PROTEIN.
FT DOMAIN 200 203 POLY-SER.
FT DOMAIN 381 384 POLY-SER.
FT DOMAIN 393 396 POLY-SER.
FT DOMAIN 493 493 POLY-SER.
FT DOMAIN 586 589 POLY-THR.
FT REPEAT 457 492 1-1.

```


PA Ball J., collection B.F.F., Kono F., Barton A.B., Su Y., Davies C.K.,
 KA Storms R.K.,
 RT "The nucleotide sequence of chromosome I from *Saccharomyces*
 KL *cerevisiae*.";
 CC J. SIMILARITY: STRONG, TO YEAST PROTEIN FILE).
 CC
 CC This SWISS-prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcements>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: 012980; AAC04471.1;
 DR SDB: S0000059; YAL064C;
 DR InterPro: IPR001389; Floculin;
 DR Pfam: PF00624; Floculin; 13;
 KW Hypothetical protein; Glycoprotein; Transmembrane; Signal;
 F1 SIGNAL 1 24
 F1 CHAIN 25 1422
 F1 TRANS-MEM 276 968
 F1 TRANS-MEM 734 775
 F1 CAP60-HYD 145 145
 F1 CAP60-HYD 167 167
 F1 CAP60-HYD 203 203
 F1 CAP60-HYD 257 257
 F1 CAP60-HYD 262 262
 F1 CAP60-HYD 276 276
 F1 CAP60-HYD 329 329
 F1 CAP60-HYD 419 419
 F1 CAP60-HYD 464 464
 F1 CAP60-HYD 509 509
 F1 CAP60-HYD 554 554
 F1 CAP60-HYD 599 599
 F1 CAP60-HYD 644 644
 F1 CAP60-HYD 689 689
 F1 CAP60-HYD 734 734
 F1 CAP60-HYD 868 868
 SQ SEQUENCE 1422 AA: 148072 MW: 148072.88; AA03011142675EA CR654;
 100% Match 3.8%; Score 131; DB 1; Length 1322;
 Best local Similarity 19.2%; Prod. No. 1.6;
 Matches 149; Conserved 92; Mismatches 28; Labels 166; Gaps 26;
 QY 42 WLRVLSQGFALPYTHPTNPRAGVTAACACGPRVQFAPGLCSAMKQVAGRA-----95
 DR 43 KNDLFTLSLEMLIVTGLNGLTDETLIVTPTATATMTTOPWNTFTLSLELTV 702
 QY 96 --- VASRRRTTAAGALTAFAHPTFVPLVDSRGALLRQYNISSPLT 144
 DR 704 IRLNGLTDETLIVTPTATATTA-MITQWNT-----FTSTSTMTVTGT 750
 QY 145 SSVAGCNLVLYAAPS -----PLPLQGNINIHIMATASNAQYVAVATIRYR 195
 DR 751 NCVPDETIVIRPTESSGLSTLTPWWTFTSTMTVTGTNSQPTDETIVILR- 808
 QY 196 PLVAVAGVYALISFEMPTTTPSVNMNSISD-----VPIIVQP--GIASELVI 246
 DR 809 IPISEGLVTTTETPRTCTIS-TSTPTTITCTNCGQDFTVIVKTPPTALSSSS 865
 QY 247 FSEGL -----HYNQWPSVETSGVAREEALISLMCLHGSLV-----N 286
 DR 866 SSGQHTSELSAKPTITTFYSNGLSVISSSDISSLSVSSSTVSSVSSPVISS 925
 QY 287 SYNIYH ICAALLLFALEERLHFNRTIVSRYSSTARHRLRGAGCTAELT 342
 DR 925 STFSSEVISTTSASHLS-PSKPSVIVPTSSFGSSPSFT-----GSSASSQS 975
 QY 343 IAAVREKMDLY---FTSTNAGVETGRG---IALLENLAUTLL-----GLPTE 386
 DR 976 SLSSESKSTSSSSSGVSAFISQELTSSLPVITIKISEQTIVTVIS-ESHCYE 1045

QY 087 LSSAGAGLVYGVVYVAVRQAFVYKSTL-----VMDAQQKQVAFPRQDL 433
 DR 1046 SSSA-----VSLAVVSSGALIE-VLWFTSLFELVSEETKQVLEETEE 1000
 QY 434 GHSRVVLDYDQHQHQRPFLPSDA-PSRPSVLRKANVMDLSLAAEYDQST 484
 DR 1052 DPLAVVLSNCSLW-SLASPAVSSSLALNGLVETLWFTSLFESKQVLEEVV 1148
 QY 485 -----YGSSTGLVYVSSVNLVNVAVGAAVAKSLIKRVILHPTSLQVYSKTF 547
 DR 1149 TSCAGVSEELSSVAVSGALATLVN-----IVAVVLSWVPEPLINE- 1190
 QY 548 VLPFAKRLKMEALIKAAVLYSTRTIAGAGLVAAVAKVAVLYVLSLQAGVPS 596
 DR 1191 --SVSSKMSNSATSETI-----NUGAAETITLGAETIKVVISSTSPNIAEQIAS 1241
 QY 597 AVAVLAHPSALALLHDI-ELC----- 1298
 DR 1242 ATDWLGHSSVSVSVSET-ELC----- 1298
 RESULT 13
 ID AMYL_SACDI STANDARD; PRO: 766 AA;
 DI 01-APR 1993 (Ref. 25; Created)
 DT 01-APR 1993 (Ref. 25; Last sequence update)
 DT 15-DEC 1998 (Ref. 49; Last annotation update)
 DE Glucanase 52 precursor (E.C.3.2.1.3) (Glucan 1,4-alpha-glucosidase)
 DE (1,4-alpha-D-glucan glucosylase) (GAL1).
 GN STAZ OR DEX1
 OS *Saccharomyces diastasicus* (Yeast)
 OC Bakaryota; Eukaryota; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos;
 OC NCBI_TaxID 41870;
 RN 111
 RP SEQUENCE FROM N.A.
 RX MEDLINE 94288586; PubMed 9017401;
 PA Kim K., Baigut J., Lee S.Y., Knudson F., Mattach L.P.;
 RT "Cloning of a new allelic variant of a *Saccharomyces diastasicus*
 RT glucanase gene and its introduction into industrial yeasts.";
 RN Appl. Biochem. Biotechnol. 44:161-185(1994).
 RN 121
 RX SEQUENCE FROM N.A.
 RX MEDLINE 91276266; PubMed 2054844;
 PA Lambrecht M.G., Proterius L.B., Schittny P., Maurer J.;
 RT "Primary structure and regulation of a glucanase encoding gene
 RT (STAZ) in *Saccharomyces diastasicus*.";
 RN Gene 100:95-103(1991).
 CC 1. CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 CC glucose residues successively from non-reducing ends of the chains
 CC with release of beta-D-glucose.
 CC 2. SIMILARITY: BELONGS TO FAMILY 15 OF GLYcosYL HYDROLASES.
 CC
 CC This SWISS Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcements>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: M60650; AAA05107.1;
 DR EMBL: M90490; AAA0560.1;
 DR EMBL: J00474; J0474;
 DR EMBL: J08017; J0817;
 DR InterPro: IPR010165; Glyco_hydro_15;
 DR Pfam: PF00723; Glyco_hydro_15; 1;
 DR PRINTS: PR00743; GLHYDRLASE1;
 DR PROSITE: PS00426; GLYC/AMYLASE_1;
 KW Hydrolase; Glycosidase; Polysaccharide degrading; Glycoprotein;
 KW Signal; Multitopic family.


```

FT CONFLICT 1234 1234 N -> Y (IN REF. 2).
FT CONFLICT 1254 1254 V -> A (IN REF. 2).
FT CONFLICT 1282 1282 G -> R (IN REF. 2).
SQ SEQUENCE 1321 AA: 2EF47F823DB9B08 CRC64;

Query Match 3.8%; Score 129.5; DB 1; Length 1321;
Best Local Similarity 22.0%; Pred.No. 2;
Matches 140, Conservative 63, Mismatches 231, Indels 201, Gaps 32;

QY 22 PQLFSSRRPRPSPSGSGSEF-----WGPVDSQP-----FAIPYIHTNPAPVPT 67
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 472 PTDMPRRPGPKINGRYQQQPEPEGLQGMEASAOPTSEAAVNQMEP--PLAMAVT 529
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 68 AAAGAGPVROPARPGLSARND---OQRPAVASRRRPTTAGAA-----PLTAVAPA 116
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 530 EMLASG-QSRSP-----WADITNEVDMPCAGS-----AGKSSPEPWLWPTNVVPS 575
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 117 ---HTTPVPVVD-SRGAILL-RQYNLSPLTSSVATGTLNLYAAPLS-----PLPL 166
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 576 ISGHSRAPVLELEKAEGPSARPATPDLFWSPLEATVSAPSPAPWEAFPVATSHDLFMAM 635
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 167 QUNTINTHIM-----ATEASNYAQYRVARATIPYPPVPNAVGGYALISFWPQITTTPT 220
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 636 LRQKFKWMLPHTPISTFANRVHAGATATAPPSPAATKV--YSLPLSL-----TPT 687
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 221 SVMNSTITINVP!!VAPGCIASPLVIPSEPLHYPNQMPSPVETSCVAEEEA?S--CLVM 278
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 688 G-----QCGEAMPTTPESPADPFETGETSPAQVNKAHSSSSPWPSPVNP 732
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 279 CIHGSIVNSYNTVYTGALGLIDFALEFRNLIPONTNTRVSRYSSTARHKLRCADGT 338
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 733 NVAVGFVPTETATEPTGLRGI-----PGSES-----GVFDT 763
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 339 AFLTTTAATRMKDLV-----FTSTNGVGEIGR-GIALT----- 371
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 764 AFSPTSGLOATDEVQNPWPSVYSKGLDASSPAPLGSPGVFLVPKVTNLPFWVAIDEG 823
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 372 -IPNIADTIGGHPTEL--TSSAGGQLFYSPVVSANGEPYKLYTSV-----ENACQK 424
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 824 PTVNPMDSVTVPAPSDASGIWEPGSQVF--EEAESTITSPQVALDTSIVTPTLTLEQSK 881
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 425 IALPHDIDIGESRVVIQDYDQHFDORPTSP-----APSRPFSVLRANDVL 471
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 882 VGVFAMSTLGSS-----SSQHPFPETQVETQCTSCASVPPHQSPLGKPAVP 929
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 472 WLSLTAAYDQSTYCSSTGPVYVSD-SVTLNVNATGAQAVARSLDWTKVTLDCRPLSTIO 530
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 940 PCTPTAASVGFSAVSSSSFTVPWDPSTLLPVTLGIE-----DPELEVLASGP----- 978
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 531 QYKTFEVLPLRGKLSWF---AGTTKA--GYPN 560
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 979 -----GVESFEEVASGEFALGTPMN 1001
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: August 5, 2003, 09:27:38
Job time : 31 secs


```

121 PVPDVSRCAILRQYNYLSPLTSSVATGNTLVLYAAPLSPLLPQDCTNTHIMATEAS 180
181 NYAQRVARATIRYRPLVPNAVGGYALISFSPWDTTTPISVIMNSITSTDVPLVQPGI 240
181 NYAQRVARATIRYRPLVPNAVGGYALISFSPWDTTTPISVIMNSITSTDVPLVQPGI 240
241 ASELVPSERLHYRNOQWRSVETSGVAEEATSGVLMCLHGSILVNSYNTNPTVGTALGL 300
241 ASELVPSERLHYRNOQWRSVETSGVAEEATSGVLMCLHGSILVNSYNTNPTVGTALGL 300
301 DFALFLEPRNLTPGNTNTRVSPYSSSTAPHRLPRGADGTAEITTTAATREMKDLYFTSNG 360
301 DFALFLEPRNLTPGNTNTRVSPYSSSTAPHRLPRGADGTAEITTTAATREMKDLYFTSNG 360
361 VGEIGRGIALTLFNLADTLGLCLTELISSAGGOLFYSRPVVSANGEPVVKLYTSVENAQ 420
361 VGEIGRGIALTLFNLADTLGLCLTELISSAGGOLFYSRPVVSANGEPVVKLYTSVENAQ 420
421 QDKGIAIHPHDIDGESKRVVIGYQDNQHFQGPHTHSPAPSPHPSVLKANDVWLSTAAEY 480
421 QDKGIAIHPHDIDGESKRVVIGYQDNQHFQGPHTHSPAPSPHPSVLKANDVWLSTAAEY 480
481 DQSTYSSSTGGVYVSDSVTLVNVATGAGAVARSLDWTKVTLGKRPSTTQQSKTFEVL 540
481 DQSTYSSSTGGVYVSDSVTLVNVATGAGAVARSLDWTKVTLGKRPSTTQQSKTFEVL 540
541 LRKGLSPWFAGTTRAGYVYNYNTTASQQLVENAAGHRVAISLYTTSIAGAPVYSISAVAV 600
541 LRKGLSPWFAGTTRAGYVYNYNTTASQQLVENAAGHRVAISLYTTSIAGAPVYSISAVAV 600
601 LAPHSALEETLIDYFARAHTEDGRTPTTPPGGAGAGTAPGSIVAHQLPKMKVKIKREL 660
601 LAPHSALEETLIDYFARAHTEDGRTPTTPPGGAGAGTAPGSIVAHQLPKMKVKIKREL 660

```

RESULT 2

```

091855 PRELIMINARY: PRI: 660 AA.
AC 091855;
DI 01-NOV-1998 (TREMREL. 08; Created)
DI 01-NOV-1998 (TREMREL. 08; Last sequence update)
DI 01-DEC-2001 (TREMREL. 19; Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=93-Egypt;
RX MEDLINE=99105430; PubMed=4890424;
RA Tsarev S.A., Binn L.N., Gomatos P.J., Athar R.F., Meillet M.K.,
RA van Ouyck-Gandre H., Tonger C.F., Indis R.L.
KT "Phylogenetic analysis of hepatitis E virus isolates from Egypt.";
R1 J. Med. Virol. 57:68-74(1999)
DR EMBL: AF051351; AAC35761.1;
DR InterPro: IPR004261; SP2;
DR Pfam: PF04014; SP2; 1;
SQ SEQUENCE 660 AA: 71040 MW: 044PF5EA7C492791 CRC94;

```

Query Match 99.18; Score 3385; DB 12; Length 660;
 Best Local Similarity 99.18; Pred. No. 5.4e-215;
 Matches 654; Conservative 3; Mismatches 0; Gaps 0;

```

QY 1 MRPRPILLLLMFLPMLPAPPDQSPSPHPSVLSGSGSPSPVVSAPFALPYRPTN 50
DB 1 MRPRPILLLLMFLPMLPAPPDQSPSPHPSVLSGSGSPSPVVSAPFALPYRPTN 50
QY 61 PRAPVNTAAACAGRPVQPARPPLCSAWPVAQWAVASRPPTTAAAILFATAVAHMT 120
DB 61 PRAPVNTAAACAGRPVQPARPPLCSAWPVAQWAVASRPPTTAAAILFATAVAHMT 120
QY 121 PVTIVDSRCATIRQYNYLSPLTSSVATGNTLVLYAAPLSPLLPQDCTNTHIMATEAS 180

```

```

121 PVPDVSRCAILRQYNYLSPLTSSVATGNTLVLYAAPLSPLLPQDCTNTHIMATEAS 180
181 NYAQRVARATIRYRPLVPNAVGGYALISFSPWDTTTPISVIMNSITSTDVPLVQPGI 240
181 NYAQRVARATIRYRPLVPNAVGGYALISFSPWDTTTPISVIMNSITSTDVPLVQPGI 240
241 ASELVPSERLHYRNOQWRSVETSGVAEEATSGVLMCLHGSILVNSYNTNPTVGTALGL 300
241 ASELVPSERLHYRNOQWRSVETSGVAEEATSGVLMCLHGSILVNSYNTNPTVGTALGL 300
301 DFALFLEPRNLTPGNTNTRVSPYSSSTAPHRLPRGADGTAEITTTAATREMKDLYFTSNG 360
301 DFALFLEPRNLTPGNTNTRVSPYSSSTAPHRLPRGADGTAEITTTAATREMKDLYFTSNG 360
361 VGEIGRGIALTLFNLADTLGLCLTELISSAGGOLFYSRPVVSANGEPVVKLYTSVENAQ 420
361 VGEIGRGIALTLFNLADTLGLCLTELISSAGGOLFYSRPVVSANGEPVVKLYTSVENAQ 420
421 QDKGIAIHPHDIDGESKRVVIGYQDNQHFQGPHTHSPAPSPHPSVLKANDVWLSTAAEY 480
421 QDKGIAIHPHDIDGESKRVVIGYQDNQHFQGPHTHSPAPSPHPSVLKANDVWLSTAAEY 480
481 DQSTYSSSTGGVYVSDSVTLVNVATGAGAVARSLDWTKVTLGKRPSTTQQSKTFEVL 540
481 DQSTYSSSTGGVYVSDSVTLVNVATGAGAVARSLDWTKVTLGKRPSTTQQSKTFEVL 540
541 LRKGLSPWFAGTTRAGYVYNYNTTASQQLVENAAGHRVAISLYTTSIAGAPVYSISAVAV 600
541 LRKGLSPWFAGTTRAGYVYNYNTTASQQLVENAAGHRVAISLYTTSIAGAPVYSISAVAV 600
601 LAPHSALEETLIDYFARAHTEDGRTPTTPPGGAGAGTAPGSIVAHQLPKMKVKIKREL 660
601 LAPHSALEETLIDYFARAHTEDGRTPTTPPGGAGAGTAPGSIVAHQLPKMKVKIKREL 660

```

RESULT 3

```

09WQAO PRELIMINARY: PRI: 660 AA.
AC 09WQAO;
DI 01-NOV-1999 (TREMREL. 12; Created)
DI 01-NOV-1999 (TREMREL. 12; Last sequence update)
DI 01-DEC-2001 (TREMREL. 19; Last annotation update)
DE ORF-2.
OS Hepatitis E virus.
OC Viruses; ssRNA positive strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AKL-90;
RX MEDLINE=99350000; PubMed=10424137;
RA Arankalle V.A., Dasanajpe S., Dasan S.D., Parrett R.B.,
RA Walimbe A.M.;
KT "Phylogenetic analysis of hepatitis E virus isolates from India (1976-1995).";
R1 J. Gen. Virol. 80:1691-1700(1999);
DR EMBL: AF124407; AAD5493.1;
DR InterPro: IPR004261; SP2;
DR Pfam: PF04014; SP2; 1;
SQ SEQUENCE 660 AA: 71941 MW: 03B72D9F0A87B521 CRC94;

```

Query Match 99.94; Score 3393; DB 12; Length 660;
 Best Local Similarity 99.28; Pred. No. 7.4e-216;
 Matches 656; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 MRPRPILLLLMFLPMLPAPPDQSPSPHPSVLSGSGSPSPVVSAPFALPYRPTN 50
DB 1 MRPRPILLLLMFLPMLPAPPDQSPSPHPSVLSGSGSPSPVVSAPFALPYRPTN 50
QY 61 PRAPVNTAAACAGRPVQPARPPLCSAWPVAQWAVASRPPTTAAAILFATAVAHMT 120
DB 61 PRAPVNTAAACAGRPVQPARPPLCSAWPVAQWAVASRPPTTAAAILFATAVAHMT 120
QY 121 PVTIVDSRCATIRQYNYLSPLTSSVATGNTLVLYAAPLSPLLPQDCTNTHIMATEAS 180

```



```

41 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520
42 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541
43 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562
44 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584
45 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606
46 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628
47 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650
48 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672
49 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694
50 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716
51 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738
52 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760
53 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782
54 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804
55 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826
56 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848
57 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870
58 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892
59 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914
60 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936
61 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958
62 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980
63 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002
64 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023
65 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045
66 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067
67 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089
68 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111
69 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133
70 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155
71 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177
72 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199
73 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221
74 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243
75 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265
76 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287
77 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309
78 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331
79 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353
80 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375
81 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397
82 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419
83 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441
84 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463
85 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485
86 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507
87 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529
88 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551
89 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573
90 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595
91 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617
92 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639
93 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661
94 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683
95 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705
96 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727
97 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749
98 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771
99 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793
100 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815
101 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837
102 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859
103 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881
104 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903
105 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925
106 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947
107 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969
108 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991
109 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013
110 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035
111 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057
112 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079
113 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101
114 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123
115 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145
116 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167
117 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189
118 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211
119 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233
120 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255
121 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277
122 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299
123 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321
124 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343
125 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365
126 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387
127 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409
128 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431
129 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453
130 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475
131 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497
132 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519
133 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541
134 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563
135 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585
136 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607
137 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629
138 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651
139 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673
140 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695
141 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717
142 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739
143 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761
144 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783
145 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805
146 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820 2821 2822 2823 2824 2825 2826 2827
147 2828 2829 2830 2831 2832 2833 2834 2835 2836 2837 2838 2839 2840 2841 2842 2843 2844 2845 2846 2847 2848 2849
148 2850 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860 2861 2862 2863 2864 2865 2866 2867 2868 2869 2870 2871
149 2872 2873 2874 2875 2876 2877 2878 2879 2880 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890 2891 2892 2893
150 2894 2895 2896 2897 2898 2899 2900 2901 2902 2903 2904 2905 2906 2907 2908 2909 2910 2911 2912 2913 2914 2915
151 2916 2917 2918 2919 2920 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932 2933 2934 2935 2936 2937
152 2938 2939 2940 2941 2942 2943 2944 2945 2946 2947 2948 2949 2950 2951 2952 2953 2954 2955 2956 2957 2958 2959
153 2960 2961 2962 2963 2964 2965 296
```


